

GenCore version 5.1.4-P5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 19, 2003, 05:12:15 ; Search time 2336 Seconds
(without alignments)
4360.438 Million cell updates/sec

Title: US-09-857-613A-28

Perfect score: 1830
Sequence: 1 MATVVRIPPTISCIHHTFRS.....IEGYKDLIKFAITCRKPE 350

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPRO/spool/US09857613/runat_12032003_100036_6013/app-query.fasta_1.519
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOFC=0 -LOOEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=humana4.cdi -LIST=45
-OUTALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=plc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09857613 @CGN 1.1.1667 @runat.12032003.100036.6013 -NCPU=6 -ICPU=3
-NO_XLPYX -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGCUT -DEV.TIMECUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vl:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vl:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rod:*
36: em.htg.mam:*
37: em.htg.vrt:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1159	63.3	1369	8 AF213481	AF213481 Perilla f
2	1154.5	63.1	1047	8 AY090280	AY090280 Arabidops
3	1154.5	63.1	1303	8 AY049258	AY049258 Arabidops
4	1151.5	62.9	1350	6 AX417728	AX417728 Sequence
5	1151.5	62.9	1350	8 AF104420	AF104420 Arabidops
6	1147.5	62.7	1296	8 AY087138	AY087138 Arabidops
7	981.5	53.6	118335	8 AC006193	AC006193 Arabidops
8	758	41.4	125422	2 AP003994	AP003994 Oryza sat
9	659	36.0	343550	1 AP003587	AP003587 Nostoc sp
10	647.5	35.4	134199	1 SYC518E	D64004 Synechocyst
11	579	31.6	108553	8 AF527809	AF527809 Sorghum b
12	416.5	22.8	298750	1 AP005375	AP005375 Thermosyn
13	392.5	21.4	334520	1 AP003588	AP003588 Nostoc sp
14	392	21.4	930	6 AX073663	AX073663 Sequence
15	392	21.4	957	6 AX073657	AX073657 Sequence
16	392	21.4	957	6 AX417730	AX417730 Sequence
17	392	21.4	974	6 AX073659	AX073659 Sequence
18	392	21.4	145709	1 D90914	D90914 Synechocyst
19	294	16.1	25681	1 SAE414559	AJ414559 Saccharot
20	287	15.7	63734	1 AF127374	AF127374 Streptomy
21	282.5	15.4	90445	1 AF040570	AF040570 Amycolato
22	270	14.8	12637	1 AF323753	AF323753 Streptomy
23	268.5	14.7	125561	8 NC866	AL513463 Neurospor
24	261	14.3	2122	6 A60301	A60301 Sequence 1
25	261	14.3	2122	6 AR144762	AR144762 Sequence
26	247	13.5	1276	8 NTU71107	NTU71107 Nicotiana t
27	244.5	13.4	5998	1 AB032524	AB032524 Streptomy
28	242	13.2	30000	6 AX250261	AX250261 Sequence
29	241	13.2	1392	8 TA060754	U60754 Trilicium de
30	237	13.0	1410	6 E07846	E07846 DNA sequenc
31	234.5	12.8	349498	1 AP003002	AP003002 Mesorhizo
32	232.5	12.7	1698	1 AF216283	AF216283 Actinopol
33	231	12.6	1629	8 AF042332	AF042332 Oryza sat
34	228	12.5	11714	1 AE012928	AE012928 Chlorobiu
35	227	12.4	1383	8 ZM079669	U79669 Zea mays en
36	227	12.4	40065	8 SPBC16E9	S299759 S.pombe chr
37	226	12.3	840	1 AF216282	AF216282 Halorhodo
38	226	12.3	1080	8 AY113031	AY113031 Arabidops
39	226	12.3	1302	8 AF375397	AF375397 Arabidops
40	226	12.3	1497	8 AF045570	AF045570 Zea mays
41	226	12.3	97554	8 AC009978	AC009978 Genomic s
42	225	12.3	6085	6 A60304	A60304 Sequence 4
43	225	12.3	6085	6 AR144763	AR144763 Sequence
44	224	12.2	1094	8 DB9131	D89131 Schizosacch
45	223.5	12.2	1267	8 NTU71108	NTU71108 Nicotiana t

RESULT 1

ALIGNMENTS

JOURNAL
REFERENCE
AUTHORS

Unpublished
2 (bases 1 to 1047)
Kim, C.J., Chen, H., Cheuk, R., Meyers, M.C., Shin, P., Ban, J.,
Bosser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D.,
Hayashizaki, Y., Ishida, J., Jones, T., Kamaya, A., Karlin-Neumann, G.,
Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,
Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T.,
Satou, M., Seki, M., Southwick, A., Tang, C.C., Tortum, M., Wu, H.C.,
Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinzaki, K., Davis, R.W.,
Theologis, A., and Ecker, J.R.

TITLE
JOURNAL

Direct Submission
Submitted (14-MAR-2002) Salk Institute Genomic Analysis Laboratory
(Signal), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL CDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamaya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y., and Shinzaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL CDNAs: Kim, C.J., Chen, H.,
Cheuk, R., Meyers, M.C., Shin, P., Ban, J., Bosser, L., Chang, E.,
Dale, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B.,
Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J.,
Quach, H.L., Southwick, A., Tang, C.C., Tortum, M., Wu, H.C.,
Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A.,
and Ecker, J.R.

Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinzaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES

source

Location/Qualifiers
1..1047
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/clone="U13225"
/note="this clone is in pU1 51
ecotype: Columbia"
1..1047
/note="gamma-tocopherol methyltransferase"
/codon_start=1
/product="At1g64970/f13011_27"
/protein_id="AA10941.1"
/db_xref="GI:19699150"
/translation="MKATLAAPSLSLTPRTNSFGSKSLFRSPSSSVSMTT
RGNVAVAATSTELARKIAEFYNTSGIMEIMGDHMHGFDSDSVSDSGHK
EAOIRMFESLPACVTDEEERKIKKVVYDVGCGSSRSYASKRGACIGITITSPV
QAKRANDLAASOLAHKASFQVADALDQFDEGKPKGLVYMSSEGEAMPKAFVKELV
RVAAPGGRITVWCHRNLSAGEALQPEONLIDKICTFTLPAVCSTDDVYNLQS
HSLQIDICADWSENAFEPWPAVIRTLWKLGLVSLRSGMKSIGKALMPLMEGYKK
GVIKFGRITCORPL"

CDS

BASE COUNT 267 a 223 c 269 g 288 t
ORIGIN
Alignment Scores:
Pred. No.: 4, 22e-92 Length: 1047
Score: 1154.50 Matches: 231
Percent Similarity: 75.438 Conservative: 30
Best Local Similarity: 66.768 Mismatches: 65
Query Match: 63.098 Indels: 21
Gaps: 4

US-09-857-613A-28 (1-350) x AY090280 (1-1047)

OY 8 Prothrllesercyslllehlsthrphearsercinsrproargthrpheala 27
DB 53 CCAAGCTCTTTCGGCTCAAGCTACGCTCTCTTT-CGGTCCATCCCTCCCTCC 111
OY 28 Argllleargval--glyproargsertrpalaprolleargalaseralaaser 46
::: ::: ||| ::::: ||| :::::

DB 112 TCAGTCTATACGACGACAAAGCGGTGGAACGTCGCTGCGCTGCTACATCCACT 171
OY 47 GUATGAGLGLUlleValleuGlulnLysProLysLysAspAspLysLysLeuGl 66
DB 172 GAG-----GCCCTAACA 183
OY 67 LysGlylleAlaGlulnPhetYrAspGlnSerSerGlyLeuTrpGlnuSnlleTrpGlyAsp 86
DB 184 AAGAGATGCGGAGGTCTACATGAATGAACCTTGGCTTCTGGGAAGAGATTGGGAGAT 243
OY 87 HisMetHisHisGlyPheTrpAspSerAspSerThrValSerLeuSerAsp-----His 104
DB 244 CATATGCATCATGAGCTTTATGACCTGATTCCTCTGTTCAACTTCTCATTCCTGCTCAC 303
OY 105 ArgAlaAlaGlulnleargmetlleGlnGlnSerLeuArpPheAlaSerValSer---Glu 123
DB 304 AAGGAGCTCAATCCGATGATGATGAGAGCTCTCTCCGTTTGCCGCTTACTGATGAA 363
OY 124 GUATGserLysTrpProLysSerlleValAspValGlyCysGlyTlleGlyLysSer 143
DB 364 GAGGAGGAGAAAAAGATAAAGAAAGTGAAGTGGTGGTGGTGGTGGTGGAGACCTCA 423
OY 144 ArgTyrLeuAlaLysLysPheGlyAlaTrpSerValGlyTlleThrLeuSerProValGln 163
DB 424 AGATATCTTGCTCTTAATTGGAGCTGAATGCATGCTGATTCCTCAGCCCTGTCAG 483
OY 164 AlaGlnArgAlaAsnAlaLeuAlaAlaAlaGlnGlyLeuAlaAspLysValSerPheGln 183
DB 484 GCCAAGAGAGCCATATATCTCGGCTGCTCAATCACTGCTCATTAAGCTTCTTCCAA 543
OY 184 ValAlaAspAlaLeuGlnGlnPropheSerAspGlyGlnPheAspLeuValTrpSerMet 203
DB 544 GTTGGCGATGCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 603
OY 204 GluSerGlyGluHisMetProAspLysAlaLysPheValGlyLeuAlaAlaValAla 223
DB 604 GAGAGCTGAGGATATGCTGCAAGGCCAAGCTTGTAAAGAGTTCGTCGTCGTCG 663
OY 224 AlaProGlyAlaLlellellellellellellellellellellellellellellellelle 243
DB 664 GCTCCAGAGCTAGCATTAATATAGTACATGCTGTCATGAATCTATCTCGGGGAG 723
OY 244 GlnSerLeuHisProTrpGlnGlnAspLeuLeuLysLysLysCysAspAlaTyrTyrLeu 263
DB 724 GAGCTTTGAGCGCGTGGAGCAAAACATCTGGACAAATCTGTAAGACCTTATCTC 783
OY 264 ProAlaTrpCysSerThrSerAspTyrValLysLeuLeuGlnSerLeuSerLeuGlnAsp 283
DB 784 CCGGCTTGGCTCCACCGCATGATTTATGTCACCTTCTCAATCCCATTTCTTCCAGGAT 843
OY 284 IleLysSerLysTrpProSerArgPheValAlaProPheTrpProAlaValIleArgSer 303
DB 844 ATTAAGGTGCGGATGCTAGAGAACGTAAGCTCTTCTGCGCTGCGGTTATACGACT 903
OY 304 AlaPheThrTrpLysGlyLeuSerSerLeuLeuSerSerGlyGlnLysTrpIleLysGly 323
DB 904 GCATTAACATGGAAGGCCCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 963
OY 324 AlaLeuAlaMetProLeuMetIleGlyGlyTyrLysLysAspLeuIleLysPheAlaIle 343
DB 964 GCATTTGACATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1023
OY 344 IleThrCysArgLysPro 349
DB 1024 ATCATTGTCGAGAGCA 1041
RESULT 3
AY049258 1303 bp mRNA linear PLN 09-AUG-2001
LOCUS AY049258
DEFINITION Arabidopsis thaliana At1g64970/f13011_27 mRNA, complete cds.
ACCESSION AY049258
VERSION AY049258.1 GI:15146233
KEYWORDS FLI_cDNA.

SOURCE
Arabidopsis thaliana.

ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi

REFERENCE
1 (bases 1 to 1303)
Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J.,
Bowler, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,
Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,
Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,
Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,
Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,
Davis, R.W., Theologis, A. and Ecker, J.R.

TITLE
Arabidopsis cDNA clones

JOURNAL
Unpublished

AUTHORS
2 (bases 1 to 1303)
Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J.,
Bowler, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,
Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,
Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,
Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,
Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,
Davis, R.W., Theologis, A. and Ecker, J.R.

COMMENT
Submitted (22-JUL-2001) Salk Institute Genomic Analysis Laboratory
(Signal), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.

TITLE
The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Cheuk, R., Chen, H.,
Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Banh, J., Bowser, L.,
Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G.,
Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,
Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A.,
Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,
Davis, R.W., Theologis, A. and Ecker, J.R.

FEATURES
Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.

SOURCE
Location/Qualifiers
1..1303
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAFL07-18-C14(R13225)"
/note="ecotype: Columbia"
1..67
68..1114
/note="gamma-tocopherol methyltransferase"
/codon_start=1
/product="At1g64970/f13011_27"
/protein_id="AAK83600.1"
/db_xref="GI:15146234"
/translation="MKATLAPSSITSLPYRNNSEFGSKSLFPSSSSSVSMTT
RGNVAAVAATSTELRKGLAEFYKETSGLMEIYMGDMHNGFTYPPDSVQSDSGHK
EAQIRMTESLRPAGYDDEEREKTKKAYDGGCGSSRLTASFGACIGITLSPV
QAKRANGLAIAOSLAHKASFDVADLDQPFEDGKFDLVSMESGDMKDAFVLELY
RVAAPGRGIIIVTCHRNLSAGEBALDQPEWDNIIDKICTEFLPAMCSTDDVNLLOS
HSLDGIKADCSERVAPFMPAVIRALTALWKGLVSLRSGMSIKALMPLMEIEYKK
GVKEFGITCKKPL"

3'UTR
BASE COUNT
ORIGIN
352 a 267 c 307 g 377 t

Alignment Scores:
Pred. No.: 5,57e-92 Length: 1303
Score: 1154.50 Matches: 221
Percent Similarity: 75.438 Conservative: 30
Best Local Similarity: 66.768 Mismatches: 65
Query Match: 63.09% Indels: 21
Gaps: 4

US-09-857-613a-28 (1-350) x AT049258 (1-1303)

QY 8 ProthrlleSerCysIleHisIleHisThrPheArgSerGlnSerProArgThrPheAla 27
DB 120 CCAACTCTTCTTCGCGCTCAAGCTACATCGCTTCTCTT-CGGCTCCATCGCTCTCC 178
QY 28 ArgIleArgVal---GlyProArgSerThrPalaProIleArgAlaSerAlaSerSer 46
DB 179 TCAAGTCTTATGACAGCAACACCGTGAACGCTGCGCGCTCTCTCATCTCCACT 238
QY 47 GluArgGlyGluIleValleuGluGlnLysProLysLysAspLysLysLeuGln 66
DB 239 GAG-----GGCTAACA 250
QY 67 LysGlyIleAlaGluPheTyraSpGluSerSerGlyLeuThrPalaunIleThrPalaSp 86
DB 251 AAAGGAATACGGAGGTTCTCATATGAACCTTCGGTTTGAGGACAGATTGGAGAGAT 310
QY 87 HisMetHisIleSerGlyPheTyraSpSerAspSerThrValSerLeuSerAsp-----His 104
DB 311 CATATGATCATATGATGCTTTATATGACCTGATCTTCTGTTCAATCTTATTTCTGCTCAC 370
QY 105 ArgAlaAlaGlnIleArgMetIleGlnGluSerLeuArgPheAlaSerValSer---Glu 123
DB 371 AAGGAAGCTAGATCCGCTATGATGATGAAGACTCTCTCCGTTTCCGCTTACTATGA 430
QY 124 GluArgSerLysThrProLysSerIleValAspValGlyCysGlyIleGlySerSer 143
DB 431 GAGAGAGAGAAAGAAAGATAAAGATGATGATGATGATGATGATGATGATGATGATGAT 490
QY 144 ArgTyrlleuAlaLysLysPheGlyAlaThrSerValGlyIleThrLeuSerProValGln 163
DB 491 AGATATCTTCCCTTAATTTGAGCTGAATGATGATGATGATGATGATGATGATGATGATG 550
QY 164 AlaGluArgAlaAsnAlaLeuAlaAlaGlnGlyLeuAlaAspLysValSerPheGln 183
DB 551 GCCAAGAGACCAATGATATCTCGCGCTGCTCAATCTCTCTCATRAGGCTTCTTCCAA 610
QY 184 ValAlaAspAlaLeuGlnGlnProPheSerAspGlyGlnPheAspLeuValThrSerMet 203
DB 611 GTTCCGATGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 670
QY 204 GluSerGlyGluHisMetProAspLysAlaLysPheValGlyGluLeuAlaArgValAla 223
DB 671 GAGAGTGTGAGATATGCTGACAGAGCCCACTTTGTAAGAGATTTGTAGTGTGCGG 730
QY 224 AlaProGlyAlaIleIleIleIleValAlaThrThrPheHisArgAspLeuGlyProAspGlu 243
DB 731 GCTCCAGAGAGTATGATATATATAGTACATGCTGCCATAGAAATCTATCTGCGGGGAG 790
QY 244 GlnSerLeuHisProThrPalaGlnAspLeuLysLysIleCysAspAlaThrIleu 263
DB 791 GAAGCTTTCAGCGCTGAGGAGCAAAACATCTTGACAAATATCTGTAGACGTTTATCTC 850
QY 264 ProAlaThrPheSerThrSerAspLysValLysLeuLeuGlnSerLeuSerLeuGlnAsp 283
DB 851 CCGGCTGTGCTCCACCATGATATATATGATGATGATGATGATGATGATGATGATGATG 910
QY 284 IleLysSerGluAspThrPheSerArgPheValAlaProPheThrProValAlaIleArgSer 303
DB 911 ATTAATGTCGGATGCTGTCAGAGACAGTACCTCTTCTGCGCTGCGGTTATACGAGCT 970
QY 304 AlaPheThrThrLysGlyLeuSerSerLeuLeuSerSerGlyGlnLysThrIleLysGly 323
DB 971 GCATTAACATGAGAGCGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1030

OY 324 AlaleuAlaMetProleuMetIleGluGlyTyrIleLysAspLeuIleLysPheAlaIle 343
 DB 1031 GCATTGACAAATGCCATTGATGATTGAAGCTTACAAAGAAAGTGCTATTAGTTGGTATC 1090
 OY 344 IleThrCysArgLysPro 349
 DB 1091 ATCACTGCGCAGACCA 1108

RESULT 4
 AX417728 1350 bp DNA linear PAT 18-JUN-2002
 LOCUS Sequence 19 from Patent WO0231173.
 DEFINITION AX417728
 ACCESSION AX417728
 VERSION AX417728.1 GI:21522874
 KEYWORDS
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
 1 Ebneth, M., Kunze, I. and Geiger, M.
 Improved method for the biosynthesis of vitamin E
 Patent: WO 0231173-A 19 18-Apr-2002;
 JOURNAL SUNGENE GMBH & CO KGAA (DE)

FEATURES
 source Location/Qualifiers
 1..1350
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 63..1109
 /note="coding for gamma-tocopherol methyltransferase"
 /codon_start=1
 /protein_id="CAD35425.1"
 /db_xref="GI:21522875"
 /translation="MKATLAASSLTSLEPTNSSFGSKSLFRSPSSSSVSMTT
 RGVAVVAATSTELRKIGAEFYNETSGLMEIDHMHGFYPPDSVOLDISGKH
 EAOIEMIEESLRFAGVDEEERKIRKVVVCGIGSGSRYSKSGAPICITLSPV
 OAKRANDLAAOSLSHKAFOYADALDOPEDGKFDLVMSGEMHMPKAEKVELY
 RYAAAGCRITIVTCHRNLSAGEPALDPRENIIDICKTFTPLPAMCSTDDVNLIS
 HSLDIDKADSENVAFTFMPAVIRALTWKGLVSLRSMKSIKALTMPLMIEGTRK
 GVIRKGIITCKPL"

BASE COUNT 374 a 278 c 314 g 384 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.07e-91 Length: 1350
 Score: 1151.50 Matches: 230
 Percent Similarity: 75.43% Conservative: 31
 Best Local Similarity: 66.47% Mismatches: 65
 Query Match: 62.92% Indels: 21
 DB: 6 Gaps: 4

US-09-857-613a-28 (1-350) x AX417728 (1-1350)

OY 8 ProThrIleSerCysIleHisIleHisThrPheArgSerGlnSerProArgThrPheAla 27
 DB 115 CCACCTCTCTTCCGCGCTCAAGAGTCAATCGCTTCTTT-CGGCTCCATCTCTCTCTCC 173
 OY 28 ArgIleArgVal---GlyProArgSerTrpAlaProIleArgAlaSerAlaIleSerSer 46
 DB 174 TCAGCTCTCATGACGACCAACGCGTGGAACGCTGCTGCGCTGCTCATATCCACT 233
 OY 47 GluArgGlyGluIleValLeuGluGlnLysProLysLysAspLysLysLeuGln 66
 DB 234 GAG-----GGCGTAA 245
 OY 67 LysGlyIleAlaGluPheTrpAspGlnSerSerGlyLeuTrpGluAsnIleTrpGlyAsp 86
 DB 246 AAGAGATACCGGAGTTTCAATGAATTCGCGTTTGCGGAGACGATTTGGGAGAT 305
 OY 87 HisMetHisIleGlyPheTrpAspSerAspSerThrValSerLeuSerAsp-----His 104
 DB 115

DB 306 CATATCATCATGCGCTTTATATGACCCGTATCTTGTGTCACTTGTGATTCGTGCAC 365
 OY 105 ArgAlaIleGlnIleArgMetIleGlnGluSerLeuArgPheAlaSerValSer---Glu 123
 DB 366 AAGCAAGCTCAGATCCGATGATTGAGAGCTCTCTCGTTTCCCGCGTGTACTGATGA 425
 OY 124 GluArgSerLysTrpProLysIleValAspValGlyCysGlyIleGlySerSer 143
 DB 426 GAGGAGAGAGAAAAGATAAAGAAAGATGAGATGTTGGGTGGATTTGAGAGACCTCA 485
 OY 144 ArgTyrLeuAlaLysLysPheGlyAlaThrSerValGlyIleThrLeuSerProValGln 163
 DB 486 AGATATCTTCCCTCTAAATTTGAGCGTGAATGCGATTTACTCTCAAGCCCTGTTCAG 545
 OY 164 AlaGluArgAlaAsnAlaLeuAlaIleAlaGlnGlyLeuAlaAspLysValSerPheGln 183
 DB 546 GCCAAGAGACCAATGATCTCGCGGTCTCAATCACTCTCTATAAAGGTTCTTCCCA 605
 OY 184 ValAlaAspAlaLeuGlnGlnProPheSerAspGlyGlnPheAspLeuValTrpSerMet 203
 DB 606 CTGGCGATCGCTTGATCGATCGCATTCGAAGATGAAGAAATTCATGATGTCGTGCATG 665
 OY 204 GluSerGlyGluHisMetProAspLysAlaLysPheValGlyLysLeuAlaArgValAla 223
 DB 666 GAGAGTGTGTAGCATATGCTGACAAAGCCAAAGCTTTGTAAGAGTTGTTGTCGTGGCG 725
 OY 224 AlaProGlyAlaIleIleIleIleValThrTrpCysHisArgAspLeuGlyProAspGlu 243
 DB 726 GCTCCAGAGAGTAGATATATATATGACATGCTGTCATGAATCATATCGCGGGGAG 785
 OY 244 GlnSerLeuHisProTrpGlnGlnAspLeuLysLysIleCysAspAlaTyrTyrLeu 263
 DB 786 GAAGCTTTCAGCGCGTGGAGCAAAACATCTTGACAAATCTGTAAGACGTTCTATCTC 845
 OY 264 ProAlaTrpCysSerThrSerAspTyrValLysLeuLeuIleSerLeuGlnAsp 283
 DB 846 CCGCTTGGTGGCTCCACCGATGATATGTAACCTGTTAAATCCATTCCTCCAGAT 905
 OY 284 IleLysSerGluAspTrpSerArgPheValAlaProPheTrpProAlaValIleArgSer 303
 DB 906 ATTAGCTGTGCGGATTTGCTGACAGACGATGCTCTTTCGCGCTGGGTTATACGACT 965
 OY 304 AlaPheThrTrpLysGlyLeuSerSerLeuLeuSerGlyGlnLysThrIleLysGly 323
 DB 966 GCATTACATGGAAGGCGCTGTCTGCTGCTGCTGATGATGAAGATATTAAGA 1025
 OY 324 AlaleuAlaMetProleuMetIleGluGlyTyrIleLysAspLeuIleLysPheAlaIle 343
 DB 1026 GCATTGACAAATGCCATTGATGATTGAAGTTACAAAGAGTGTCAATTAGTTGGTATC 1085
 OY 344 IleThrCysArgLysPro 349
 DB 1086 ATCACTGCGCAGACCA 1103

RESULT 5
 AF104220 1350 bp mRNA linear PLN 05-JAN-1999
 LOCUS AF104220
 DEFINITION Arabidopsis thaliana gamma-tocopherol methyltransferase (g-TM)
 mRNA, nuclear gene encoding plastid protein, complete cds.
 ACCESSION AF104220
 VERSION AF104220.1 GI:4106537
 KEYWORDS
 SOURCE Arabidopsis thaliana.
 ORGANISM Arabidopsis thaliana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
 1 (bases 1 to 1350)
 Shintani, D.K. and DellaPenna, D.
 Elevating Vitamin E content of plants through metabolic engineering
 Science (1998) In press
 JOURNAL Shintani, D.K. and DellaPenna, D.
 REFERENCE

TITLE Direct Submission
JOURNAL Submitted (03-NOV-1998) Department of Biochemistry, University of Nevada, Reno, NV 89557, USA
FEATURES Location/Qualifiers
source 1. 1350
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
gene 1. 1350
/gene="g-TMT"
63. 1109
CDS
/gene="g-TMT"
/codon_start=1
/product="gamma-tocopherol methyltransferase"
/protein_id="AAD02882.1"
/db_xref="GI:4106538"
/translation="MKATLAAPSLTSLPRTNNSPFGSKSLIFRSPSSSSVMTT
RGNVAATASTFALRGIKAEFNETSGTCAEITDHHHFFYPDDSVQSDGKH
EAOIRMIRESLFAGYATDEEEKIKKVDVCGGSGSRVYASKFGACIGITLSPV
QAKRANDLAAQSLSHKASFOVALDOPEDGKFDLVWSMESGEMDKAFVELY
RYVAPGRITIVTWCNHRMSAGEALQPEWQNLIDKIKTFYLPAMCSTDYVNLDS
HSLDIKCADMSENYAPFWPAVIRALTKGLVSLRSGMSIKALTMPLMEGKK
GVIKGIIITQCPV"

BASE COUNT 374 a 278 c 314 g 384 t
ORIGIN

Alignment Scores:
Pred. No.: 1.07e-91 Length: 1350
Score: 1151.50 Matches: 230
Percent Similarity: 75.43% Conservative: 31
Best Local Similarity: 66.47% Mismatches: 65
Query Match: 62.92% Indels: 21
DB: 8 Gaps: 4

US-09-857-613A-28 (1-350) x AF104220 (1-1350)
QY 8 ProthHILserCysIleHISLHISLTHrPheArGSeRGLnSerProArGThrPheAla 27
DB 115 CCAACTCTTCTTGGCGTCAMAAGTCAGCGTCTCTTT-CGGCTCCATCCTCCTCC 173
QY 28 ArgIleArgVal---GlyProArGSeRTPAlaProIleArgLaseRAlaIaseRser 46
DB 174 TCAGTCTCATGACGCAACCGCTGAAACGCTGCGGCTGCTGCTACATCCACT 233
QY 47 GluArGlyGluIleValLeuGluGlnGlyProLysLysAspLysLysLeuGln 66
DB 234 GAG-----GGGCTAAGA 245
QY 67 LysGlyIleAlaGluPheTyrsAspGluSerSerGlyLeuTriPGLuAsnIleTriPGLyAsp 86
DB 246 AAGGATATAGCGGAGTCTTACATGAACACTTCGGCTTGGGAGAGATTGGGAGAT 305
QY 87 HISMETHISHisGlyPheTyrsAspSerThrValSerLeuSerAsp-----His 104
DB 306 CATATCCATCATGCGCTTTATATGACCTGATCTTCTGTTCAACTTCTGATTTCTGCTC 365
QY 105 ArgAlaAlaGlnIleArGnetIleGlnGluSerLeuArGpHeAlaSerValSer---Glu 123
DB 366 AAGGAGCTCAGATCCGCTGATGAAGAGTCTCCGTTCCCGCTGTACTGATGA 425
QY 124 GluArGSerLysTriProLysSerIleValAspValGlyCysGlyIleGlyGlySerSer 143
DB 426 GAGGAGAGAAAAAGATAAAGATGATGATGCGGTGGGATTTGGAGAGAACTCA 485
QY 144 ArgTyLeuAlaLysLysPheGlyAlaThrSerValGlyTleHrLeuSerProValGln 163
DB 486 AGATATCTTGCCCTTAATAATTGGAGCTGAATGCATTGCACTCACTCAGCCCTGTTCA 545
QY 164 AlaGlnArGAlaAsnAlaLeuAlaAlaGlnGlyLeuAlaAspLysValSerPheGln 183
DB 546 GCCAAGAGAGCCATGATCTCGCGCTGCTCATCACTCTCTCATGAAGCTTCTTCCAA 605
QY 184 ValAlaAspAlaLeuGlnGlnProPheSerAspGlyGlnPheAspLeuValTriPserMet 203

DB 606 GTTCCGAGATCGTTGCATCAGCCATTCGAAGATGAATAATTCATCTAGTGTGCGATG 665
QY 204 GluSerGlyGlnHISMetProAspLysAlaLysPheValGlyLeuAlaArgValAla 223
DB 666 GAGAGTGATGACATATGCTGACAGGCCAAGTTTATAAAGGTTGATGACGCGG 725
QY 224 AlaProGlyAlaIleIleIleIleValThrTriPcysHisArGAspLeuGlyProAspGlu 243
DB 726 GCTCCAGAGAGATGAATAATATAGTACATATGCGCCATGCAATATCATATCGCGGGAG 785
QY 244 GlnSerLeuHisProTriPGLuGlnAspLeuLeuLysLysIleCysAspAlaTyTriLeu 263
DB 786 GAAGCTTTGGACCGCTGGGAGCAAAACATCTTGACAAAATCTGTAGACGCTTATCTC 845
QY 264 ProAlaTriPcysSerThrAspTyValLysLeuGlnInsLeuSerLeuGlnAsp 283
DB 846 CCGGCTTGAGCTCACCGCATGATATGATCACTGCTTAATCCATTCCTCCAGAT 905
QY 284 IleLysSerGluAspTriPserArGpHeAlaAlaProPheTriProAlaValIleArgSer 303
DB 906 ATTAAGTGTGCGATTTGTCAGAGAAGCTAGCTCTTTCGGCTGCGTTATACGACT 965
QY 304 AlaPheThrTriPcysGlyLeuSerSerLeuLeuSerSerGlyGlnLysThrIleLysGly 323
DB 966 GCATTAACATGAGAGGCGCTTGTCTCTCTCTCTGCTGATGATGAAGATATAAGCA 1025
QY 324 AlaLeuAlaMetProLeuMetIleGluGlyTyrIleLysLysAspLeuIleLysPheAlaIle 343
DB 1026 GCATTAACATGCGCATTAATGATTAAGGTTTACAGAAGAGTGTCAATTAAGTTGATC 1085
QY 344 IleThrCysArgLysPro 349
DB 1086 ATCATTGTCAGAGCA 1103

RESULT 6
AY087138 1296 bp mRNA linear PLN 26-JUN-2002
LOCUS Arabidopsis thaliana clone 32143 mRNA, complete sequence.
DEFINITION AY087138
ACCESSION AY087138.1 GI:21405862
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1296)
Haas,B.J., Volkovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
Full-length messenger RNA sequences greatly improve genome
annotation
Genome Biol. (2002) In press
2 (bases 1 to 1296)
Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
Full-length cDNA from Arabidopsis thaliana
unpublished
3 (bases 1 to 1296)
Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
Direct Submission
Submitted (11-MAR-2002) Ceres, Inc. 3007 Malibu Canyon Road,
Malibu, CA 90265, USA

COMMENT
TITLE
JOURNAL
REFERENCE
AUTHORS
This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the

C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the WS or Ilaer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Geneset carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

FEATURES

source

1..1296

/organism="Arabidopsis thaliana"

/db_xref="taxon:3702"

/clone="32143"

68..1114

/codon_start=1

/product="gamma-tocopherol methyltransferase"

/protein_id="AA06496.1"

/db_xref="GI:21592747"

/translation="MKATLAAPSSLTSLPYRTNSSEKSSLLFRSPSSSSVMTT

RGNVAVAATATSTALRKIAEFYNETSGLEWEDHMHGFPDSSVOLSDGKH

EAOIRMEESLRFAGVDEEBEKIKKYVDCGCGSSRYLASKFGAICITLSPV

OAKRANDLAASLSHRSKASFOVADALDPEDEKDFDVMSESGHMDPKAFVKEIV

RYAARGRIITVTMCHRLNSAGEALQMEONILDKICTFPLPAMCSTDDVYNLOS

HSLDDIKCADMSENVAPFPVIRATLTKGLVSLRSGMSIKGALIMPLMEIGYK

GVIRFGITTKCPK"

BASE COUNT 352 a 269 c 303 g 372 t

ORIGIN

Alignment Scores:

Pred. No.: 2,27e-91 Length: 1296
Score: 1147.50 Matches: 229
Percent Similarity: 75.43% Conservative: 32
Best Local Similarity: 66.18% Mismatches: 65
Query Match: 62.70% Indels: 21
DB: 8 Gaps: 4

US-09-857-613a-28 (1-350) x AY087138 (1-1296)

QY 8 ProthrllesercysllehlsilhslnrPheargserGlnserProargThrPheala 27
DB 120 CCACTCTCTTGGCGCTCAAGTCAATCGCTTCTTT-CGGCTCCATCTCTCTCTC 178
QY 28 ArgileargVal--GlyProargSerThrPalaProileargAlaSerAlaSerSer 46
DB 179 TCAGTCTATGACGACACACGCGTGAACGCTGTGCGCTGCTGCTACATCCACT 238
QY 47 GluargGlyGluIleValleuGlnGlnProlysLysaspLysLysleuGln 66
DB 239 GAG-----GCGCTAAGA 250
QY 67 LysGlyIlealagluPheThrArgSerGlyLeuThrPgluasnileTrrPglYASP 86
DB 251 AAGGAATACCGGAGTTCACATGAAACTTGGGTTTGGAAGACGATTGGGGAGAT 310
QY 87 HismethHisHsIglyPheThrArgSerSerThrValSerleuSerAsp-----His 104
DB 311 CATATGCATCATGCGCTTTTATGACCCGATTCTTCTGCTCACTTCTGATTCGCTCAC 370
QY 105 ArgAlaAlaGlnIleargmetIleGlnGlnSerleuArgPheAlaSerValSer--Glu 123
DB 371 AAGGAGCTTCAGATCCGATGATGAAGACTCTCTCGTTTCCCGGCTTACTGATGAA 430
QY 124 GluArgSerLysThrProlysSerIleValAspValGlyCysGlyIleGlyLysSer 143
DB 431 GAGCAGCAGAGAAAGATAAAGAAAGTAGTGTGGGTGGGATGGAGAGCTCA 490
QY 144 ArgTyrLeuAlaLysLysPheGlyAlaThrSerValGlyIleThrLeuSerProValGln 163
DB 491 AGATATCTTCCCTTAATTTGGAGCGCATGATTCGATTCATCTCAGCCCTGTCAG 550
QY 164 AlaGlnArgAlaAsnAlaLeuAlaAlaGlnGlyLeuAlaAspLysValSerPheGln 183
DB 551 GCCAAGAGACCAATGATCTCCGCGCTGCTCAATCACTCTCTCATTAAGGCTCTTCAA 610
QY 184 ValAlaAspAlaLeuGlnGlnProPheSerAspGlyGlnPheAspLeuValTrrSerMet 203

DB 611 GTTGGCGATCGCTTGATATGACCATTCGAAAGATGGAATAATTCATCTATGCTGCGAG 670
QY 204 GluSerGlyGlnHisMerProAspLysAlaLysPheValGlyGluLeuAlaArgValAla 223
DB 671 GAGAGGTGATCATATGCTGCTGACAGGCGCAAGTTGTAAAGATGGTACTGCTGGCC 730
QY 224 AlaProGlyAlaIleIleIleIleValAlaThrTrpCysHisArgAspLeuGlyProAspGlu 243
DB 731 GCTCCAGAGGTAGATATATATAGCATGATGGTGCATAGAAATCTATCTCGGGGAG 790
QY 244 GlnSerLeuHisProThrPgluGlnAspLeuLysLysIleCysAspAlaTyrTyrLeu 263
DB 791 GAAGCTTTCACCCCGTGGAGGAGCAAAACATCTTGACAAATCTGTAAGACGTTCTATCTC 850
QY 264 ProAlaTrpCysSerThrSerAspTyrValLysLeuGlnGlnSerLeuSerLeuGlnAsp 283
DB 851 CCGGCTTGTGCTCCACCGATGATATGTCACATTCCTCAATCCATTCCTTCACAGGAT 910
QY 284 IleLysSerGluAspTrpSerArgPheValAlaProPheThrProAlaValIleArgSer 303
DB 911 ATTAAGTGTGCGATTTGTCAGAGACGATGATCTTCTGCGCTCGCTGATACGAGCT 970
QY 304 AlaPheThrTrpLysGlyLeuSerSerLeuLeuSerSerGlyGlnLysThrIleLysGly 323
DB 971 GCATTAACATGAGAGGCGCTTGTCTGCTGCTGCTGATGATGAAAGTATTAAGA 1030
QY 324 AlaLeuAlaMerProLeuMetIleGlnGlyTyrLysLysAspLeuLysPheAlaIle 343
DB 1031 GCATTGACATGCTGATGATGATGAAGTTGACAAAGAGCTGCTATTAAGTTGATTC 1090
QY 344 IleThrCysArgLysPro 349
DB 1091 ATCACTTCCAGAGAGCA 1108

RESULT 7

AC006193/c

LOCUS

DEFINITION

AC006193

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

TITILE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITILE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITILE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITILE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITILE

JOURNAL

118335 bp DNA linear PLN 28-MAY-2000
Arabidopsis thaliana chromosome I BAC F13011 genomic sequence,
complete sequence.
AC006193
AC006193.3 GI:473953
HTG.
Arabidopsis thaliana.
Arabidopsis thaliana.
Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta,
Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots,
Rosidae, eurosids II, Brassicales, Brassicaceae, Arabidopsis.
1 (bases 1 to 118335)
Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,
Altati, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P.,
Gonzalez, A., Kremetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S.,
Luros, S., Schwartz, J., Shinn, P., Toriumi, M., Vysotskaia, V.,
Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.
Unpublished
2 (bases 1 to 118335)
Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,
Altati, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P.,
Gonzalez, A., Kremetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S.,
Luros, S., Schwartz, J., Shinn, P., Toriumi, M., Vysotskaia, V.,
Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.
Submitted (09-DEC-1998) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
3 (bases 1 to 118335)
Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,
Altati, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P.,
Gonzalez, A., Kremetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S.,
Luros, S., Schwartz, J., Shinn, P., Toriumi, M., Vysotskaia, V.,
Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.
Direct Submission
Submitted (04-MAY-1999) DNA Sequencing and Technology Center,
JOURNAL

Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
4 (bases 1 to 118335)
FedeSPIEL,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altati,H., Araujo,R., Huizar,L., Kowley,D., Buehler,E., Dunn,P., Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S., Luoto,S., Schwartz,J., Shinn,P., Toriumi,M., Yotsukaia,V., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (11-JUN-1999) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
5 (bases 1 to 118335)
FedeSPIEL,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altati,H., Nguyen,M., Lam,B., Southwick,A., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (28-MAY-2000) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
On May 4, 1999 this sequence version replaced gi:4678192.
COMMENT
FEATURES
Source
1..118335
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="1"
/map="q1147"
/clone="F13011"
1092..2311
/gene="F13011.1"
join(1092..1225,1474..2311)
/gene="F13011.1"
/note="Hypothetical protein"
/codon_start=1
/protein_id="AAD38245.1"
/db_xref="GI:5042406"
/translation="MGQLIDPCKTNTQNEWMFIESEKLSAFQRTAQTIVKLOO
RRPTKMKDKMOSPEMTATRIPPDPSTGVEDAMSVCKPKLVNPNYSCSS
OOHOKRAKAFINDINFIKYMESELCADDEIKELAEIDYERKARRAEIMIKLAK
DYEEEMAREAEEMOKRLFKELSEKSEKSVKRRKDLSEEROMHRLAEVLEERVVK
LMDARLLEKSELEENRQGERENRMMKPKILLERACSSPARRCEPQIKRGINP
PPRVRAIRSKSEKWKLECKVQKILRKQTPRCPTPLSPSP"
complement(3352..3963)
/gene="F13011.2"
complement(3352..3963)
/gene="F13011.2"
/note="Hypothetical protein"
/codon_start=1
/protein_id="AAD38246.1"
/db_xref="GI:5042407"
/translation="MGNCIFGGLGDEEDLLIKYKSDGCVLEFISPVTAGVSHGFS
GHALFSVDLMLKPLAHDLVPGOSYLPNIVSDELKTFVGSCHVRNSSELSAT
PYRMSLDYHNRVLKRSYDVFSRNSHIRTROKERTRRRTSSKGAIMKVNLIITEE
LQILSEDRTELLEISYRAVAKGETSSITSSSENFSLSVQT"
8107..9734
/gene="F13011.3"
join(8107..8140,8222..8358,8440..8489,8562..8893,
8959..9062,9144..9219,9304..9734)
/gene="F13011.3"
/note="very similar to alcohol dehydrogenase. Similar to
formaldehyde dehydrogenase. [Arabidopsis thaliana];
Location of est 94M15T7 (g11275/072) and 94M15X (3.)
(g113450565)"
/codon_start=1
/protein_id="AAD38247.1"
/db_xref="GI:5042408"
/translation="METQKVTTCRAAVAKGCEPLVMEDVYVDPQRLERVIRILFT
SICHIDLSMKENENQRAIPRIILHAGAGIVESGVEEMAGGQHVLPITGEGD
CVRCKRDGANLIERFRVPMKRVMTDGTGTRFTSDNKPPIYHPLNTSTFSEYVTL
ACVLKVDPLEPIKISLSCGVSTCKLMNEGVAMNADVADPASTVAIFGLAVGL
AVAEGARARGASKIIGIDINPDKFQJGREAGISEPINFKESDKAVHERVMEITEGVE

YSPFCAGSIEALREAFSTNSGVGVYVAAVSPOLPIPHMETLFGSRITASVGC
FKRQLPPEPTQCLGLDILDLFISHQLEPFIHINEMQMLHOGKRLRGLHL"
complement(10308..11958)
/gene="F13011.4"
complement(join(10308..10729,10802..10883,11042..11125,
11208..11388,11461..11617,11723..11958))
/gene="F13011.4"
/note="Location of ests 66H377 (g112414136), 62B10T7
(g11931264), 228J15T7A (g111217083). 89% identical to
membrane related protein Cp5 [Arabidopsis thaliana]"
/codon_start=1
/product="putative membrane related protein"
/protein_id="AAD38248.1"
/db_xref="GI:5042409"
/translation="MNGLSLVTFEFSHSCFSDLIYFAPRLMVAAGLVGVWVRPK
VADLSDLSNSPKFLNQLPTSLTKTSSHVWVSLSEKTEGVDDDFRLMKLVE
VKDGGCWLOMDMRSTPTPSYQAMRDPKPGPQSRPVEFADPEMDRLMKLVE
RSKMDMLYSSTLERCKDGTIVGVWVWVKRPFKFCSDREYIISRLIMDAGRYVCITK
GVQYPSVPKPRRVDLYSSWCIRAVSGKSPSAGTMCATMATHNPKVSAEPEMNGSIAE
LVROGMCVAKKRIEPLGLRAYORAKAGCLSPSAGTMCATMATHNPKVSAEPEMNGSIAE
VGDRTGKNIRKILYVGAIALACTLDKGLTKAVIFGVARRFRMRKM"
13535..15717
/gene="F13011.5"
join(13535..13647,13808..14022,14535..14734,14807..15315
15402..15717)
/gene="F13011.5"
/note="Identical to Tubulin Alpha-1 Chain [Arabidopsis
thaliana] (g11135391) Location of ests 139J20T7
(g112762968), 214C14T7 (g11159108)."
/codon_start=1
/product="alpha tubulin"
/protein_id="AAD38249.1"
/db_xref="GI:5042410"
/translation="MREIISIHGAGCIQVNSCMELCYLCHICIQPDGTPMSDSTVGA
CHDAENTFEFSHSCGVHRAVFLDEPVIDEVRGTROLFHPQLISGEDAANN
FARGHYTVREIYDCLERKLADNCQLOGLFVNAAGVSGSLLERLSD
FGKSKLGTFTYSPQVSTAVVEYVSMTSLHLETTVVVLDNLEALYDLCRSID
IEPPYSNLRILISOTISLTSLRFDGAINDIPEFQINIVPYPRIHMLSSYAPI
SSAKAHDEPSPVETITSVSPESNMMAKDDPRHKGIMACIMRGQVYKRVNTAVAA
IKAKRTIYVDMCPGFGKGINIQPSVPGDGLAVQVACMISNNTVAEVSID
HKFDLMYSRAVHYVGGMEGESEARDELALALEKYEVEVGCGADDDEEGDEY"
complement(16208..17133)
/gene="F13011.6"
complement(join(16208..16338,17040..17133))
/gene="F13011.6"
/note="Unknown protein; Location of ests 247F21T7
(g112413341)"
/codon_start=1
/protein_id="AAD38250.1"
/db_xref="GI:5042411"
/translation="MAAEPRKATVAEYKMDLEDEDEFEFEINEDMLEKEEYKVSQ
QWEDWDVDDVDNDFSRQLRKELENGTDRK"
complement(17641..19522)
/gene="F13011.7"
complement(join(17641..18251,18334..18833,19281..19522))
/gene="F13011.7"
/note="31% identical to glucan endo-1,3-beta-glucosidase
precursor ((1-3)-beta glucan endohydrolase) [Triticum
aestivum] (g111706551)"
/codon_start=1
/product="Similar to glucan endo-1,3-beta-glucosidase
precursor"
/protein_id="AAD38251.1"
/db_xref="GI:5042412"
/translation="MATRKLPEKTVVQMLKNNINKVKLEFDEDTMGALAGGLEVM
VAIPNDQLKVMISYDRAKDWKRVNTRVAFDGVNITPAVNEPELISYNGSFNLT
PALANIONALNDAEGNSVAKATVPLNADVYSPASNPVPSGRRPDITIGOMTOYD
FLKNNAPITITINIPPLSLTGNDPEPLNLAFFDGAEPINDNGIDITNVFADANFDLVS
SLKAVNGCMPLIIVGEWPTGGRHANAGAVRYNGILPLPLGINKTKPLRPYIEV
YLFGLDEPAKSIAGCPFERHMGIFKFDQGPFPIDLSGCGSKFLIAQVAVPYLPNK
WCTFPEARDELTLAANDIYACTFSDCTALFQDAGNSVYFAFNNFQVKNQ
DESACYFQGLATITITONISQGCNRPDIQIVASASSFSSSVSLVILVWFLSLGMMF


```

QY 122 ----- 122
Db 16914 TAGTACTGCAAGAAATAGCAAAAGACCAAAATTGTCCTGATTCAGCCGTACTTTAC 16855
QY 122 ----- 122
Db 16854 CCTTGGAGTACGTACATGCTCAATTATTTGCTCTACGTATATATTTGCTTGATGCT 16795
QY 123 ---gluGluarGserLysTrpProLysSerIleValAspValGlyCysGlyIleGlyGly 141
Db 16794 GCACATGATGCGGAGAGAAACCAAAAGTGTAGTTGATGTTGGCTGTGGCATTTGGTGT 16735
QY 142 SerSerArgTyrIleuAlaLysLysPheGlyAlaTrpSerValGlyIleThrLeuSerPro 161
Db 16734 AGCTCAAGATACCTTGGCGCAAAATAGCGAGCGCAATGCTACGGCATCAGCTTGATGCG 16675
QY 162 ValGlnAlaGlnArgAlaAsnAlaLeuAlaAlaGlnGlyLeuAlaAspLys----- 179
Db 16674 GTGCGAGCTGAAAGAGAAATGCCCTCGCGAGAGCAAGGCTATTCAGACAAAGCTGCGT 16615
QY 179 ----- 179
Db 16614 ATTACTACTGTTTATCTGTTCTAATAAAATTCTACTGTTTATTCGATCGGATTTAA 16555
QY 179 ----- 179
Db 16554 TCTGCGTGTACTGATCATATTACATTGCAATATCAGAGGTTTAAAGCTGTGATTAC 16495
QY 180 ----- 180
Db 16494 TCACCTGCTGTGTGACGCAATCTGTTTAAAGAGCTCTCTCTTCAAGTTGGTGATGCAT 16435
QY 188 uGlnGlnProPheSerAspGlyGlnPheAspLeuValTrpSerMetGluSerGlyGlnI 208
Db 16434 GGACAGAGCTTTCTCTGATGGGACGTTGATCTGTGTGTCATGAGAGTGGGAGCA 16375
QY 208 smetProAspLys----- 212
Db 16374 CATGCGAGACAAAGCGGAGTAAATACTCCTCTTTTATCTTACAGAAAAAAGAAA 16315
QY 212 ----- 212
Db 16314 ATGGGAAATGTAAATCCCTCCGTTTACAGTTATAAGACTTCTTACATTCGCCAGTTCA 16255
QY 212 ----- 212
Db 16254 ATACATGTTAATGAATCAGCACACATATATGTTTACATTCAATATATATGAAT 16195
QY 212 ----- 212
Db 16194 ATGGGTATACTAGAAAGTTTATATATATGAAGGAAAGAAATATCTTGATCGGAGAAC 16135
QY 212 ----- 212
Db 16134 GTTTCAAATCAGCCATCAGGTTAAGTGTGATGGTTACCATGATCGTTGATCTT 16075
QY 212 ----- 212
Db 16074 GGTGAGCAATCCATCTGTGATTTAATTTCCGGGGTGGGAAAAATGTAGAAAAACGA 16015
QY 212 ----- 212
Db 16014 GAGGATATTATGAACAATCTGCTATTTATTAATAATCAGATGCAATCAATCATCTT 15955
QY 212 ----- 212
Db 15954 GAGAGAGTCTGTATTGATGTTACTGATAATGTTTGGATCTTTTCACTCTCTTTT 15895
QY 213 ----- 213
Db 15894 AGCTAAGAGTTCTTACCTGAGTCTTTTACAGTTTACAGGAGCTGGCAGCGCTCCAC 15835
QY 224 apProGlyAlaIleIleIleIleValThrTrpCysHisArgAspLeuGlyProAspGlu 244

```

```

Db 15834 TCCTGGGCGGAGAAATTAATCATTTGACCTGCTGCATGAGAACTCGAGCATCGAAGA 15775
QY 244 nSerLeuHisProTrpGluGlnAspLeuLeuLysLysIleGlyAspAlaIleTrpLeuPr 264
Db 15774 GTCCCTGAACCTGTATGCTGATGCTGATCTGATAAAGATATGCGATGATATATCTCC 15715
QY 264 oAlaTrpCysSerTrpHisArgPyrValLysLeuLeuInsLeuSerLeu----- 281
Db 15714 AGACTGTGCTCTCTCTTCTGATTTATGTCATAAATTGCCAGTCAGCTCTCTTGAGTAA 15655
QY 281 ----- 281
Db 15654 AAACTTTTCATGCTCTGAACTCGTAAGTGAATTTAAAGTACAACTGATATGTTTGA 15595
QY 281 ----- 281
Db 15594 CATCACTTGGCTGATACATCCGATTTGCAATTCCTCCAGACATCTTGCATGTGTGAC 15535
QY 282 ----- 282
Db 15534 ATGTGAATGTGCCAGATATATAGACAGCTGATTTGTCAGAGAACGTGCGCCATTTCTG 15475
QY 297 pProAlaValIleArgSerAlaPheThrTrpLysGlyLeuSerSerLeuSerSerG 317
Db 15474 GCCTCGGTTATTAATATCAGATTCATGATGAGAAAGTTTAACTTCTCTGTAAGTGTG 15415
QY 317 yGln----- 318
Db 15414 TATGATCTTGGCCATCTCTCTTCTCCACTATATGATTTACGCAAAACAGATCTTGACA 15355
QY 319 ----- 319
Db 15354 AACGTAACTAATTTGTGCTGCTGCTTAAATTGAAGCGTGGAAGACGATTAAGAGT 15295
QY 324 lAlaLeuAlaMetProLeuMetIleGluGlyTrpLysLysAspLeuIleLysPheAlaIle 344
Db 15294 CAATGTGATGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 15235
QY 344 lEthrCysArgLysProGlu 350
Db 15234 TCACCTGTGCAAGCCCGAA 15215

```

```

RESULT 9
LOCUS AP003587 343550 bp DNA linear BCT 28-NOV-2001
DEFINITION Nostoc sp. PCC 7120 DNA, complete genome, section 7/19.
ACCESSION AP003587 BA000019
VERSION AP003587.1 GI:17130808
KEYWORDS
SOURCE Nostoc sp. PCC 7120 DNA.
ORGANISM Nostoc sp. PCC 7120
Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
REFERENCE
1
AUTHORS Kaneo, T., Nakamura, Y., Wolk, C.P., Kuritz, J., Sasamoto, S.,
Watanabe, A., Iriuguchi, M., Ishikawa, A., Kawashima, K., Kimura, T.,
Kishida, Y., Kohara, M., Matsumoto, M., Matsuno, A., Muraki, A.,
Nakazaki, N., Shimo, S., Sugimoto, M., Takazawa, M., Yanada, M.,
Yasuda, M., and Tabata, S.
TITLE Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120
JOURNAL DNA Res. 8 (5), 205-213 (2001)
MEDLINE 21595285
REFERENCE 2 (bases 1 to 343550)
AUTHORS Kaneo, T.
DIRECT SUBMISSION Direct Submission
INSTITUTE Submitted (02-MAY-2001) Takakazu Kaneo, Kazusa DNA Research
Institute, The First Laboratory for Plant Gene Research, Yana
1532-3, Kisarazu, Chiba 292-0812, Japan
(F-mail: kaneok@kazusa.or.jp,
URL: http://www.kazusa.or.jp/cyanobase/,
Tel: 81-438-52-3935 (ex. 2338), Fax: 81-438-52-3934)
FEATURES
Location/Qualifiers

```


source 1. .343550
/organism="Nostoc sp. PCC 7120"
/db_xref="taxon:103690"
/note="synonym: Anabaena sp. PCC7120"
gene complement(94. .411)
/gene="a111719"
complement(94. .411)
/gene="a111719"
/note="ORF_ID:a111719
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB73418.1"
/db_xref="GI:17130809"
/translation="MKKSAILEAVHETAEPLHKAAGLMNQTTLKEFEHLCPRIEPL
VPLQIKETRESSQSOAVFARILINISTSVQKWEIGQKRPASLKLHLVKNRGLNS
VLY"
gene complement(504. .782)
/gene="a111720"
complement(504. .782)
/gene="a111720"
/note="ORF_ID:a11720
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB73419.1"
/db_xref="GI:17130810"
/translation="MKIKYKRTFEDMARREGKLSLCAVNEAPGLYDADIGGLF
KRIKAPGCKSGSGFRTLIATNNEVGFLFAFKTNNAVTLTHMKRKL"
gene .1507
/gene="a1r1721"
935. .1507
/note="ORF_ID:a1r1721"
/codon_start=1
/transl_table=11
/product="transcriptional regulator"
/protein_id="BAB73420.1"
/db_xref="GI:17130811"
/translation="MRHKDKKKQAICDAIELITANGPADTSMKIAKANSPATI
VYFENKEKLNLKILYFVKOEISAETIKGVNOMLSEKAKMTWNNYKKAIPNPAE
LAFEOFANSPMVRICKDESLSYFKPMLDLFNGKKEKVKFDPLEIFTAFTAPLNG
LKEHFGSLIIDEKMLETTYKIAMAVTN"
gene .2504
/gene="a1r1722"
1569. .2504
/note="ORF_ID:a1r1722"
probable oxidoreductase"
/codon_start=1
/transl_table=11
/protein_id="BAB73421.1"
/db_xref="GI:17130812"
/translation="MKKEKMAENITIPKQGLRIAVTSSSGIGYETARVLANKOASYI
IAYRNDCKKAKALAKITLOQKNDKDYKMEIDLNLASVKNPAEMFRKNYLRDLINN
AGVAPIPSKTTGCFELQFCTNHLGHALGQLEFISTBGSRIIVVSSAHNMGTI
DPPDLNMEORSYAKRWAKYGSKLANILFYTELQRLKLDNGIDTLVYASHPQWTELO
RTAGGIYKVLNGIYAODITMGALPTLRAALEGLKGAIEYGPNGFMEMRGYPIVEEN
ELSKDOLAKKLWVSEKLTDVKEPFMKKQASAG"
gene complement(2701. .3462)
/gene="a111723"
complement(2701. .3462)
/gene="a111723"
/note="ORF_ID:a11723
probable aldolase"
/codon_start=1
/transl_table=11
/protein_id="BAB73422.1"
/db_xref="GI:17130813"
/translation="MSYFSRPLPPVFNSTIEBERLHRKORLAARLFGKFGESGIAG
HITRDEFTDFHWNPGLKYEFGHIRVSDLVNRDGVVGDALINDAARAIHSQIH
EARDVVAHAHASYGKAMSLGRILDPLODACAFYEDHALDDYGVAVLETSEK
RLAETLGNKKALILRNHILVGHVTEAARWYISLERSCOALLAEAAQPKTIIGHE

gene complement(3764. .5167)
/gene="a111724"
complement(3764. .5167)
/gene="a111724"
/note="ORF_ID:a111724
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB73423.1"
/db_xref="GI:17130814"
/translation="MSTSQGVPGCLDPLTHPELIDLGSLITASLLVLYIAGIA
IALMSFALRHNSQOPIFTGEWAIRYAOILRGLOHALVLLVGGFISLTSNRTH
HMEQAKIAQVVOVSQDKLEQIAPQVRYTOEPTVNTVNGKIVYNDTQOVSRLLS
LGSIOVLKEOSTDVQGRSAVYRADYADVKNRDKDINNEFEIOPNGSYLLSN
KYKERORTLEPVPNGVSPAFLEPGOETSPFVYVQAQGPWVYNSAGOLSNFRL
TIANFPGADFASGIYPERKADGRSTQVTHIFDDNVSVKNPGEVPTNTPIRONGIT
PRLLLAPALEFWMLILLVLSPLISLTVNAISITFACILITLTIYSRMDAOLAMT
ISLVLLTWGLGSTRSASLAIITCIAGAVVVLGLLVPSGLTSLVAGLSVIMLA
VRHVMNMINWOPEDSKS"
gene complement(5492. .5902)
/gene="a111725"
complement(5492. .5902)
/gene="a111725"
/note="ORF_ID:a111725
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB73424.1"
/db_xref="GI:17130815"
/translation="MFLKPPSGLMVNASLTLNCAITVYSGRSIAEAKSNFPLYQNGA
DORGFPIALDASINGTFTLVQDHGIGIAKRRLHAACSGRLFTYKFSLYAANGKIT
SEDKTEOEIFPEPKRTPDASMEIYCVANNHOKK"
gene .6842
/gene="a1r1726"
6321. .6842
/note="ORF_ID:a1r1726"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="BAB73425.1"
/db_xref="GI:17130816"
/translation="MKSQHLQATTLAGEVSQYLKEERTALNANAELODFIDMCPDARE
VRKLAVKIYQVGLYDEIOKIIDPSSGTGMKQAAVEEGIDRLRYNGRKSYLENA
KQDOEVLSTWOTKCMELGELYOLAEIYDVYVESKSYIDLPSEACISMKTKYVP
KADENAVAEKKRD"
gene .7431
/gene="a1r1727"
6919. .7431
/note="ORF_ID:a1r1727"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="BAB73426.1"
/db_xref="GI:17130817"
/translation="MMGDLSSGYWGSQDOETAVPVNPDQKOTYGAADVLLGELYLK
AYDAGNSKNTNLYLEYLIANSPPDRLITFWGASIHNSKEIRGLDSVNSLPTBQK
IHCVFAPNCPVONDIEDIWLQARTWRFICALIPSFHLKWIEMFLRHTTFDPTL
OMGYYSKIK"
gene .9401
/gene="a1r1728"
8064. .9401
/gene="a1r1728"
/note="ORF_ID:a1r1728
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB73427.1"
/db_xref="GI:17130818"
/translation="MTGKNARNRNAPLRLVSGNGAFAFGSESRYSITSKEVYIGRPSCQ
VLDDAMYRMVSRRAVVRPVAVSSVDSKFSWVLCDLNSANGTYINGRLGCGOELHAG

gene 986. .1933
/gene="slr0073"
CDS 986. .1933
/gene="slr0073"
/note="ORF_ID:slr0073"
/codon_start=1
/transl_table=11
/product="sensory transduction histidine kinase"
/protein_id="BAA10540.1"
/db_xref="GI:1001703"
/translation="MGEMPSPKEGTTLDQOKILGFIIEAOPHELTLEKGLIDLCYV
VDSSEVQNEFRASHIKGGAALGTTSTOKTAHRLDEDAKVLKEHQLPVDQLESIF
LNGYVDLQDLVERKIQSPAGIQPEADSIYGAAYOEELQNTLYNLNOSARATSTQ
TAAAKSTGTGTENSDVPEIKVLHKMLQLEKOEPSATSRKLEQECYCTGRAPKO
ENMHLSTLSKQAIANPLHSYRTLAIPVVLKDKQGYDCFLAKPAGNIRVSDGLQGLAS
SOVAOYLIPADPAASAVLMKVPFNDSQLRLKRLNSAAR"
complement(1930. .2652)
/gene="sl10088"
complement(1930. .2652)
/gene="sl10088"
/note="ORF_ID:sl10088
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAA10541.1"
/db_xref="GI:1001704"
/translation="MGVLSVYWTATHRPQNFLLLLTMTLSSSHSTKEDIILCYLLKE
GGSAIAMAEEGIFSPQARKHLKDLKENDGLIEHODEROSMGRPOFLYQISKGRQF
PQRGEFALSTDSLVEYVGEEDGAVLKQOMORKKAEATROQIGGCPLSRYKRLVEL
RKQEGMAEIHPLSVEQAEKFIJSEHHCALADVAESYPTVCGHELMFAAILPDCATE
RTHMLDNGEHTGYLIQSKTPN"
2908. .4350
/gene="ycf24"
2908. .4350
/note="ORF_ID:slr0074"
/codon_start=1
/transl_table=11
/product="ABC transporter subunit"
/protein_id="BAA10542.1"
/db_xref="GI:1001705"
/translation="MSSTVKNLVNOPYYYGYVTNIADAIPIRGLSEDEVRLISAKKN
EPEFMDFLRLAYRHWLTMAEPTPAVHPPIIDYODIIVYSAPKSKKLESIDEDP
ALLETTEKLGIFSEQKRLSNVAIDAFDSVSGTTEKKEIAEDGVIPICSIQLQEH
PLVQKYLGSVPEADNPFALNSAVESDGSFVFIPKVCPEMLSTYRINNGDGO
PFRLLIIAEGASVSYLECTAPMTVDNOLHAAYVELVALDNADIKSYVQNNYADDE
NKGCIYVTRKGLCKGVNSKISWTQVETGSAITTKYKSCVLYCNGNSYGEFISALT
NNKQADTGTAKMHIIGKNTRSIIISKIGISAGNSANSYRGLVKNKGPRAQAGARNYSQDS
MLIGRAAANTPEPIYIOVDNNTAKVEHEASTSKIGEDQLFYFAORGISEBDAAVSMLVSG
PCKDVLNELPMPEFAAEADKILSLKEGTG"
4492. .5262
/gene="ycf16"
4492. .5262
/note="ORF_ID:slr0075"
/codon_start=1
/transl_table=11
/product="ABC transporter subunit"
/protein_id="BAA10543.1"
/db_xref="GI:1001706"
/translation="MSQTLISIKNLASVADGNOILKGVNLEIKAGEVHAIMGNGSGK
STLSVITGHPYEITNGEIIITYGODLSALEPHERALAGIFLAFOYPLPEPGSNIDF
LRIAVNAKKHKLGELEDTDFEDLIDQELDVVAKMNPALERSLNGEFGSGEKKREI
LQMTLLEPTLSLIDELDSGLDADALRIYSEGVNPLKNPNATLVITTHYORLLNYIIPD
HIIHWYDKIVWSGGKEALBLEEKGYDPLDQVLAPI"
5398. .6759
/gene="slr0076"
5398. .6759
/note="ORF_ID:slr0076
hypothetical protein"
/codon_start=1
/transl_table=11

gene /protein_id="BAA10544.1"
/db_xref="GI:1001707"
/translation="MTAAIILADLITQKHQTEIDKELOAKKEQAMVSGTSDSLVKNR
OGAADLGSRLDPKDEBEQFTDSELAKIDVAAKKSVDVAAAEFTLPEAHQSR
LVIFNGEFTPELSTNDLPBSAITQOSMTNLAHQRLAVLYQKTDGNEFENLNAH
GMTDSAVVMIPANTELKSPIHLLEFLTVYDPTLMVOPRLVYVYENNAQVYIAGSYGAI
STNCTDRPOQOPEYFNIVSEIYLGENAQVYHINORSGSGFVIATATIAQKOSRR
LIDVNLGAKLSRNLQMTQOEEATKTEFTALLTLAGQOSQVATIALNPHGATNKL
HKCIVDEYQAQVRSKGVNDQAQQLTNAOQLNNVILVLSARINTREPELOITADNKC
SHGATTSQLEADEVFYLRSGNDYDRHLLIDAFAGEIIDDQIPLASVQRLKQCVSC
RTI"
6827. .8089
/gene="nifs"
6827. .8089
/gene="nifs"
/note="ORF_ID:slr0077"
/codon_start=1
/transl_table=11
/product="Nifs protein"
/protein_id="BAA10545.1"
/db_xref="GI:1001708"
/translation="WVALDIPSLAATVRODEPIILNDEINCHPLVYLDNANTSOXPRAY
LEKLMHYENDNANVARGAHLQLSVRATDAVEARNVYAKFINARSREIYTTNATEA
INLVAYSMGMMNLKAGDEIITTYMEHNSMLVPQMYAAKGAALKFVQLEQESFLE
HEFTLSEKTLVTVYHISNTLGCVPABEILAOALHAQAKKVLVDAQSAAPHYLDVQ
LIDCDMLVASGHMKCAPGTGPIYKGEELIENAMPREPGCEGMAIEVFPDFTGELPH
KREAGPTAIEAIALGAAYDITLDCMENTHNYEVELTHYLMQGLQITPOLRLGPNP
KHODRALASFVYAGLASDVATVMDGCIATRSGHCTQPLRLPFDASGSARALYF
YNTKEIDLEFLOSQATIRFSDDEFTV"
complement(8079. .9269)
/gene="sl10086"
complement(8079. .9269)
/gene="sl10086"
/note="ORF_ID:sl10086
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAA10546.1"
/db_xref="GI:1001709"
/translation="MRVILMTGKGGVKTSVVAATGRLCAELCHKTLVUSTDPAHSLA
DSFDLELGEHPRLVKENLGAELDALMEGNMGAVKRYITQVLAQGLDGVAAEELA
ILPGMDEIFGLVPMKRRHYDEADYDLIIDSAPYGTALRLSLPEVGMWYRRYKPIQ
GMSVALRPLEVPRPIAGFSLPDKQVMDAPYEFQIEALEKVLNDNOTSVRLTN
PEKMYIKESLRAHYSLVNSYDVLITANRIIPETIDDPFORWKINOQVYKQEIYDN
FHLPIPKAEKSEEMKCGLAALERLKTDTLKXDDPSQVYIKENTINIVQSSNDYSIE
LVLPKPKQIDLNKGTGDELVNRITGNHRRNLVLPQALALSPAGAMMEDYLRIRAE
AVKR"
complement(9312. .10523)
/gene="sl10085"
complement(9312. .10523)
/gene="sl10085"
/note="ORF_ID:sl10085
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAA10547.1"
/db_xref="GI:1001710"
/translation="MNSDHSROQPLFISLIPHLMGKGHIVPYHGAVRSVEYLQMY
QAMVPQDADPNLEISDRRLILPEDLAEGSMVAKITRLPAMQCSVDYLAQQLQSL
LNGVESRPVITLLEFRIHLQLAIMMGLQLSPEARSRLHFWILYRDDVLDQTSIY
HMLHORLRLKPLPPQCOPLTDSPPGLRSLODYEGOPTVPIHPHOFAEVYPLFTEDR
PLTCMWPAGAPREKSGSVMRHIVSLPRSPRSQQLTAAASGLKSSGSEIEVALIGDR
LTREERAGLAKAQLIILPYDEPVAERTSGITTEAITERITPLVSPGTWMARELQY
DIDSLVDNQAOOSLVWAEIDOKLAOSSIHQOLAAMAQAVRRHHSLESYAKTIOETWLS
RANGNNEVNNM"
complement(10520. .11005)
/gene="hisB"
complement(10520. .11005)
/gene="hisB"
/note="ORF_ID:sl10084"
/codon_start=1
/transl_table=11
/product="imidazoleglycerol-phosphate dehydratase:"

histidinol-phosphatase"
/protein_id="BA010548.1"
/db_xref="GI:1001711"

Alignment Scores:

Pred. No.: 5,43e-45 Length: 134199
Score: 647.50 Matches: 135
Percent Similarity: 63.16% Conservative: 45
Best Local Similarity: 47.37% Mismatches: 102
Query Match: 35.38% Indels: 3
DB: 1 Gaps: 3

US-09-857-613a-28 (1-350) x SYCSLRF (1-134199)

QY 65 LeuGlnLysGlyIleAlaGluPheTyrAspGluSerSerGlyLeuTrpGlnIleTrp 84
DB 27990 CTTCTAGCAAAAATTTAAATTTCTAGCAGCAGCTCCAGCGCTCTGGAGACGTTTGG 28049
QY 85 GLYASPHISMeHisHisGlyPheTyrAspSerSerThrValSerLeuSerAspHis 104
DB 28050 GGTGACATATGACACACCGCTACTACGCTCCACGGCACCCTATGGGATC---GATCGC 28106
QY 105 ArgAlaIleAlaGlnIleArgMetIleGlnLysSerLeuArgPheAlaSerValSerGlu 124
DB 28107 CGCGACGCTCAATGATGATCAAAAGAACTATGCGCTGGCA---GTGCCCAAAAT 28163
QY 125 ArgSerLysTrpProLysSerIleValAspValGlyCysGlyIleGlyGlySerSerArg 144
DB 28164 AGCGCCAAA---CCACGAAAATTCGATTTAGCTGTCGCAATGCGGAGCTAGTTTG 28220
QY 145 TyrLeuAlaLysLysPheGlyIleThrSerValIleThrLeuSerProValGlnAla 164
DB 28221 TACTTGGCCAGACACACGAGAGAGATGAGGGGCTGATCTTCCACAGTCAGTG 28280
QY 165 GlnArgAlaAsnAlaLeuAlaAlaGlnGlyLeuAlaAspLysValSerPheGlnVal 184
DB 28281 GAACGGGCGGGGAGAAAGGCGCAGGCTGGGCTGGCTCAACCTCCAGTTTCAGGTG 28340
QY 185 AlaAspAlaLeuGlnGlnIlePheSerAspGlyGlnPheAspLeuValTrpSerMetGlu 204
DB 28341 GCCAATGCCCTTGATTTGGCTTCTCCGATTCCTTACGTGGGTTGCTGCTGGAA 28400
QY 205 SerGlyLysHisMetProAspLysAlaLysPheValGlyLeuLeuAlaArgValAlaAla 224
DB 28401 AGTGGGAGACATGCCCAACAAAGCTGATTTTACAAAGCTTGGCGGACTTAA 28460
QY 225 ProGlyAlaIleIleIleIleValThrTrpCysHisArgAspLeuGlyProAspGluGln 244
DB 28461 CCAAGTGGCGCTGATTTAGGACCTGCTGTCATCTCCATTTGATCCCGCAATGGC 28520
QY 245 SerLeuHisProTrpGluGlnAspLeuLysIleCysAspAlaTrpTyrLeuPro 264
DB 28521 CCCCTGCTCCGATGAACCTCCGATCCCAACCCATCATGACGTTTACTGTTTCCC 28580
QY 265 AlaTrpLysSerThrSerAspTyrValLysLeuGlnSerLeuSerLeuGlnAspIle 284
DB 28581 TATGTGCTTTCCCTCCCGGACATGAGCGCATCCCGAGGAATGTGGGGAATTT 28640
QY 285 LysSerLysAspTrpSerArgPheValAlaProPheTrpProAlaValIleArgSerAla 304
DB 28641 AAGACTCCGATGTGATGAGGGGCGTCACTTTTGGGACCGCGGATGATGATGCGC 28700
QY 305 PheThrTrpLysGlyLeuSerSerLeuLeuSerSerGlyLysThrIleLysGlyAla 324
DB 28701 TTGCTATCCCGGCTGTGGGCTTGGGCAACCGCGGCGCAAAATATTCATGATGCGCC 28760
QY 325 LeuAlaMetProLeuMetIleGlnGlyTyrLysLysAspLeuIleLysPheAlaIleIle 344
DB 28761 CTGCTTTAGCATTAAGAAATGGGCTATGAGGGGATTAGTGCCTTTGGCTTATTA 28820
QY 345 ThrCysArgLysPro 349
DB 28821 ACGGGGATTAAGCTT 28835

RESULT 11
AF527809/c
LOCUS
DEFINITION
Sorghum bicolor clone BAC SB_BBC0234M12 pnp200725 orthologous region.
ACCESSION
AF527809
VERSION
AF527809.1
KEYWORDS
GI:22208503
SOURCE
sorghum.
ORGANISM
Sorghum bicolor
Spermatophyta: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Eukaryota: Magnoliophyta: Liliopsida: Poales: Poaceae: PACC clade: Panicoidae: Andropogoneae: Sorghum.
1 (bases 1 to 108553)
Song,R., Liaca,V. and Messing,J.
Mosaic Organization of Orthologous Sequences in Grass Genomes
2 (bases 1 to 108553)
Song,R. and Messing,J.
Unpublished
REFERENCE
TITLE
Submitted (08-JUL-2002) The Plant Genome Initiative at Rutgers - Waksman Institute, Rutgers University, 150 Frelinghuysen Road, Piscataway, NJ 08873, USA
JOURNAL
FEATURES
source
1. 108553
/organism="Sorghum bicolor"
/cultivar="BTx623"
/db_xref="taxon:4558"
/chromosome="G"
/clone="BAC SB_BBC0234M12"
/note="overlaps with sorghum BAC SB_BBC0126P21"
1. 108553
/note="pnp200725 orthologous region"
complement(4347..>7716)
/gene="SB234M12.1"
/note="similar to Triticum aestivum Yr10 gene GB:AF149112.1 GI:11990496; similar to Hordeum vulgare M1a1 gene GB:AY009939 GI:11612212"
/evidence=not_experimental
complement(join(4347..4451,4587..4676,4917..4990,5038..6794,6896..>7716))
/gene="SB234M12.1"
/product="putative stripe rust resistance protein Yr10"
complement(join(4347..4451,4587..4676,4917..4990,5038..6794,6896..7716))
/gene="SB234M12.1"
/note="putative 108 kDa protein"
/codon_start=1
/evidence=not_experimental
/product="putative stripe rust resistance protein Yr10"
/protein_id="AA094319.1"
/db_xref="GI:22208504"
/translation="MEVVTGVLTVGLFKLKLTVGEYNLQKGVGEIMFLQPELQSKQ GALKESKAPSDLDHODRIMASEVELSYDIDSIDMEFYRREGGQULAAADGKMG FIDRSIDLTRFRRVROVAMDIDIRRVAREAREAYKIDGVGAPVDPVPRLLA HTAYVEIINGIDARDELKIVLTDDGSOEAKOHGVAEISVGGIGKTLAVVOK IRQDFCMAFVSOTPDNRRLFEGLISELGRKINETRYRRFIDALIKFLQTRKC ITVIDD1WD1SVKMICALPDNNGGVITTTTFNFAVEISGAYSKALCHESRKL FTYRIGNEEKYKCPDEHLVESHRLKNCAGPLA1ITIASLANKARAKMELEV NSIGTGEDSTVDENMK ILAYSVDLKYRLVCLLISWEPEDYPTKNNHLLMMWIA EGFVOCBOGKSLFELGECYENELISMDIOPVDRHEAMIEHC RVDMVLEVRISLN EDNEVITLNNHSTSSQKFRRLSLONSVHLASPPRSVMQVSVITFSSALQOM PALASRVLVNLKGSYLPQGDHLKHLGULPHURITGLDRTTNELPREICNLRILQ TLDOVGSYISLPSTVAVQDHTLMLCIVDQMTVPNGIGRLTALEUSLTVTCESTRI PEBELCHTELRLVLEFMMNTLEKSLVACCKLRKIOSLIVASGEYVENDAVAPRI VORLOOSCFSESLPDMNNSPLGLSLPIANVRELROEDLETGRPALNHLVLAKD H EKLIRRFRTVGCSPFCIVGCLGFMGAVFPGAMIKLRSLSLTLVARDVREIMA SSDVGCVIITFPSCGANEMEVAKAAVAYSTTEITENTLNCIDQOKKMMNMCKE IHTDELKRRNNMGIDTRGSKNSYVPELIRCK"
/rpt_unit="99a
8776..8802
complement(412094..>14263)
repeat_region
gene

```

/gene="SB234M12.3"
/note="highly similar to Oryza sativa gene for putative
receptor kinase GB:AF128457 GI:4680335; similar to Oryza
sativa receptor kinase-like protein gene GB:AF119222
GI:4680488; similar to Arabidopsis thaliana receptor-like
protein kinase emb:AL138654 GI:689879"
/evidence=not_experimental
complement(join(<12094..12725,12823..14263))
/gene="SB234M12.3"
/product="putative receptor kinase"
complement(join(12094..12725,12823..14263))
/gene="SB234M12.3"
/codon_start=1
/evidence=not_experimental
/product="putative receptor kinase"
/protein_id="AAM94320.1"
/db_xref="GI:22208505"
/translation="MYTLAFLRLPEFLLLAAAAAANGNNTTEEAAMALVNL
KSFPTPTGRLKMSATSPFPCDASPFGVOCYKSLVGRLTTHNLSTGDFGAI
AKLRLHSLVNLKHNALSGPLPASLGLRLGLALYLSNNFSGPIPADVFANMLKTL
YLDNNRITGCLPADAIKAPRLMELHLDNRQIDGPVFKLPESLRNVSNNRLKTL
POGVAERYDASSFAGNPGLCSPGSDAAYCAAGPALPAPPTADYAMEESTSV
FVIGIILYLLVSGAMVLMRLDDENSVATAMDYAGSTTGAGAGAGSGSKSA
DAASTAARACGAVAVDMASSGSSSHGCGSSSRBRGERTLLMEDIPTAGLPDL
MKASAEYLKGTGSAKKAAMRGVTVAVRRLDMNRFGRDEEQHVHMLGQLRHPV
LPVGHYLRKEKLIYSEYPRGSLYLHGDSPNRYLLDMOGRVRIAGVYRGALF
LHRLGIPARLVSMDGADDPAPPPHGNLCSGNILLGADMEPLLVDFGEPFLVA
AQAPQAMFARSEGTGRGVYASVDYCGYGLVLETVGRPSOYLDSRGCDVYH
MAATVAAGEAEVLDPALIAAGDAAVRLRLRGVHVASPEPCRSVALAAMVERI
GSAAGAS"
repeat_region
12558..12578
/rpt_unit-c99
complement(13145..13307)
/rpt_unit-c99
13490..13517
/note="GC-rich"
14190..14216
/rpt_unit-c99
complement(<15173..>20442)
/gene="SB234M12.4"
/note="similar to Arabidopsis thaliana gene for putative
sugar transporter protein GB:AY094465 GI:20453188"
/evidence=not_experimental
complement(join(<15173..15605,16042..16226,16382..17055,
17854..18423,18531..18650,19447..19501,19596..19649,
20293..20442))
/gene="SB234M12.4"
/product="putative sugar transporter"
complement(join(15173..15605,16042..16226,16382..17055,
17854..18423,18531..18650,19447..19501,19596..19649,
20293..20442))
/gene="SB234M12.4"
/note="AT-rich"
/codon_start=1
/evidence=not_experimental
/product="putative sugar transporter"
/protein_id="AAM94321.1"
/db_xref="GI:22208506"
/translation="MSMICKLNVQADMLMGATGAGQGVOKNGCKLEAVNLGCAH
RIMAAVSTYHOVCAMTFRFACATASRLIDELATREFSGFTIVCKRAIVLDMK
STVSAVVVSGIYALMDWPALEALNHNHKKFELLNGPSIEGITLAASTFGAIVIT
IFSGLLDMLGRRAILVYSLVSGVLMWLPNIVYLALIGSGSGLVETVP
IYISETSPNNRGLGTMPQMFICTIFSICLIPWLLMSSNNHMLIGAIIPASIV
YFALYYLLESPRMLASDGKISEARVSLORLGRKHDSGETAIVYDCVDTISDAV
GTARASFSGTASRTWSPSTFWOISDPLVDLGSIHENMSGGSRRNSFFVFNSF
SPREHNNHBRGNSDOOTREATISAGEVNNGLASLISQAAVEVDANSTPSE
GSSYLRHNGTSVLAQEFMASTHDHDEEBEELHGLVPHQSAPRMESGRHOVRL
SETADSKFMKRVLLPGIRHALCYGMILALQOSTISGLRCAPEILRQVAGLESD
IGSPHSTSLISTLQALLPCTITAMLMDCGRVIVLATPTLLISLWSMSI
LFNMGPDAIVHPALTYOCFSYVIGLCPPIPIKLSLSEPTKARATCASFCSLSWF
GGLLSAYCLPVMKSTIGLGCAGCIYALVCCAPLFLFYTRIPETKMLLELAEFLS
RQRYV"
repeat_region
complement(22125..22161)

```

```

repeat_region
/note="AT-rich"
complement(22424..22446)
/note="AT-rich"
repeat_region
23019..23089
/note="AT-rich"
repeat_region
23130..23172
/rpt_unit=ta
23495..23526
/note="AT-rich"
repeat_region
26812..26900
/note="AT-rich"
repeat_region
37014..37038
/note="GC-rich"
complement(37399..37462)
/rpt_unit=a
<37578..>44288
/gene="SB234M12.6"
/note="similar to Nicotiana plumbaginifolia ubp1 gene
emb:A172011 GI:696559"
/evidence=not_experimental
join(<37578..37721,38745..38806,38998..39100,40241..40317,
40518..40579,40690..40777,40913..40988,41080..41110,
41205..41318,42137..42222,42364..42537,43066..43258,
43335..43537,43864..43887,43980..>44288)
/gene="SB234M12.6"
/product="putative oligouridyate binding protein"
join(37578..37721,38745..38806,38998..39100,40241..40317,
40518..40579,40690..40777,40913..40988,41080..41110,
41205..41318,42137..42222,42364..42537,43066..43258,
43335..43537,43864..43887,43980..44288)
/gene="SB234M12.6"
/codon_start=1
/evidence=not_experimental
/product="putative oligouridyate binding protein"
/protein_id="AAM94322.1"
/db_xref="GI:22208507"
/translation="MHPRMKAQAADQOLMQLALLQOOAOOPLEPGHHHPGL
AAPMEPIVSGNLPFGPDSSTGSRVYGNISLOTDVIVLOEVSIGPVGCKLIRE
KSSRFDTDYDRRYAALATLSLNGRPLXGPITVYNNATSTQEDISGHNITVDIC
PEYTDATLPAFSGSICSDARPMMDKTRSGFGVSRNOODASALINDNGWL
GNQIRCNMAITKGANAGEERQIVDSKVDLNGTSESGKNPNDEGSENOFTTVVG
NLPEHATINDVHLFPHSLGAGSIEYAVTRDKGFGFVRYSTHEAAQAIOMANGOLIG
GROIKTTIYNGVYKRLAVEGSDCEVLYLWSRLTVLASLVVAVFLLEGDAYRO
GIDLLAASVAGAVGCGVAVGVITVYVAOSDPEGLITMKKACACACVIRR
SKCGSVKCAKGAQVAVRYVTENLGNKSKSPPTPTASSPLPPAPALFSTGV
SPDPLAIONMALNKATNPISLGMQALNOVALMGGRASQSTYDGGFQGISSTAATT
SAQQOQOQOQOLMY"
repeat_region
37580..37668
/rpt_unit-cag
complement(39121..39142)
/note="AT-rich"
44247..44274
/rpt_unit-cag
complement(45252..45289)
/rpt_unit-c99
<45290..>46757
/gene="SB234M12.7"

```

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
4,17e-39	108553	579,00	160	38,908	61	242
Percent Similarity:		38,908				
Best Local Similarity:		32,598				
Query Match:		31,648				
DB:		8				

US-09-857-613a-28 (1-350) x AF527809 (1-108553)

OY 94 AspSerAspSerThrValSerLeuSerAspHisArgAlaIleGlnIleArgMetIleGln 113

DB 94288 AATTACACTAGTATATAAATTGGTTGAC-----ATCGTCCCACTTGA 94241

OY 114 GluSerLeuArgPheAlaSerValSerGluGluArgSerLysTrpProLysSerIleVal 133

gene
CDS
complement(1482. .2213)
/gene="ho2"
/note="ho2"
/codon_start=1
/transl_table=1
/product="heme oxygenase 2"
/protein_id="BAC09273.1"
/db_xref="GI:22295447"
/translation="MTVLSLARREGTAEAHTLAHTAFMKEFVRGLITSRLRPROFLAN
LYEVYDLEALASANSADAVRAIFYPELINSQGLAODLAFYGEDMRDQVPTLATRY
YLSRLHDVAOSEPLLLAHSTYRYMGDSGOALSK IVHSOLMLPEKGRAPFEPAL
PTMEDKAFKORATLDSGLDEEALRIYQEANFAPALNCNMHMLSESLKTTID
QVMOSWIENQPRROHREPALAI"
2486. .3550
/gene="tlr11722"
2486. .3550
/gene="tlr11722"
/note="ORF_ID:tlr11722"
hypothetical protein"
/codon_start=1
/transl_table=1
/protein_id="BAC09274.1"
/db_xref="GI:22295448"
/translation="NVAIAFNPHSGGQSVKTIKRENLTLPRFTTDFEQAKLIDSRQOE
OLEAMLEMRADYNRRHFEVRDPSFKKIDNOTKRAFDIYERSCVSEPSGFLRK
ELSRKLDKRRPLAEIFHLMARDBARHAGLNKAMDFGAMIDGYLSKRTYTFEPL
PWLVTVYSEKIGYWKYIIIVHLEKHPEHSFYPLFNFERMCOENRIGDIFKALI
HSQPOLIOGAKLMHNFLLTVAFHTLVHERAKFEALGLDAPAFDEVEIARTNE
TARTFSVLANVDHPEFSYRLQRCVEISNQLQIIEATHQKMKALKRPHQALIAH
LRIIVLPVNAQCEWGTVM"
complement(3547. .3858)
/gene="tlr11723"
complement(3547. .3858)
/gene="tlr11723"
/note="ORF_ID:tlr11723"
hypothetical protein"
/codon_start=1
/transl_table=1
/protein_id="BAC09275.1"
/db_xref="GI:22295449"
/translation="MLSRPLIERLAAEIGRDYIMDAKWHLYIGDAKLATPLAEAF
PLLEQEMSLAQTSLVQEVSVPLGGDQQLSLERLIPKANQELVVELNRFQEELE
S"
gene
CDS
complement(3963. .4382)
/gene="psad"
complement(3963. .4382)
/gene="psad"
/note="ORF_ID:tlr11724"
/codon_start=1
/transl_table=1
/product="photosystem I subunit II"
/protein_id="BAC09276.1"
/db_xref="GI:22295450"
/translation="MTTLTGQPLVGGSTGGLISADTEEKYAITMSPKEQVEMPT
AGAAVMRBEENLVYPRKKEQCLALAQQLRPRIINDYKIRIPDEETVLIHKPDGV
PEKVNKGRVAVNSVPSIGQNPNSQLKFTGKKPYD"
4763. .5398
/gene="tlr11725"
4763. .5398
/gene="tlr11725"
/note="ORF_ID:tlr11725"
hypothetical protein"
/codon_start=1
/transl_table=1
/protein_id="BAC09277.1"
/db_xref="GI:22295451"
/translation="MOYLYLHGFASSPRSSKAOYFFDREBELCQPLIIDLNOEDFEH
LTLROIQVEALTLPEEPTITLIGSSFGGLTAMWLAEKHPVVQLEFLAFAFEFAIHM

gene
CDS
complement(5383. .6369)
/gene="tlr11726"
complement(5383. .6369)
/gene="tlr11726"
/note="ORF_ID:tlr11726"
/codon_start=1
/transl_table=1
/product="delta(24)-sterol C-methyltransferase"
/protein_id="BAC09278.1"
/db_xref="GI:22295452"
/translation="MSHGLIVTVVAVALVGLALYLEPFRKESARSVAESYDMNT
KDGLIEFWGCHILGHYGLPPKDFROAKVDVHEMVMRAGLDRLPCPTTVYDVC
GIGSSRLIADGFEHVTGIPISPEQVRRARELTPRAELNRFQDDALASFPSASD
VYMSIEGSRPHNDKQOQAKELLIRLKGCGILVADMRORDROOPLWFEELIRQL
DQNAHPAPASIEGFBEALATGLVAGVMTADMTQETLPSMLDSIWGIVAPCEGLIRF
GLPLVKSLEPVPFLMRIAFCMGLCRFGMFAVRVRAEIPAVSLERAPQVNC"
6619. .8262
/gene="rbaA"
6619. .8262
/gene="rbaA"
/note="ORF_ID:tlr11727"
/codon_start=1
/transl_table=1
/product="riboflavin biosynthesis protein Riba includes
GTP cyclohydrolase II and 3,4-dihydroxy-2-butanone
4-phosphate (DHBP) synthase"
/protein_id="BAC09279.1"
/db_xref="GI:22295453"
/translation="MPVFDSTIESALALRGEVIVVVDENRENEGDLGAERVPRA
MINMVAHARCLILAMEGRDLDELPLVYTTNTDSNOTAFYVSVDAGRMCVTTCI
SAEBRARTIALIDPSODPLRPRGHVPLRSRPGVILRAGHTEAAYDLTRAGLY
PAGYICIQNDGMSRLPOLMEARTHQIKTISADLISRLQHEFERVREVAKIP
TERGEFOITGRNSLDQSEHAIVKGDPAFEGQPVIVRHSCELTGDALGSLRCOR
MQDLAALKMINAAGGVVYLRQEGRGIGVNLKRAYSLQDLQDVEANEHNGFPAD
LRNYGVAQIILNDGKQIRLITNNPKIAGLGVLEVDVRYPLLEATPYNTPLYT
TKAEKLGHLILQTYLITVAIFRMQSGSLAADREKLEKLRHLAAGVHLIAREAPVA
OALRGHDLVHFGEDQPHVADRWQYKOPHPVYLAVQTLIRQFCQPLTKGEGFLVA
TGDPMLNLPMDDLRQATQOOSPSBMRPHILYMSAAG"
complement(8267. .8453)
/gene="ts11728"
complement(8267. .8453)
/gene="ts11728"
/note="ORF_ID:ts11728"
unknown protein"
/codon_start=1
/transl_table=1
/protein_id="BAC09280.1"
/db_xref="GI:22295454"
/translation="MDVRTDKQCHNSVIVYKSLAPLNKRVTTRPEKCVGNLGVVP
TPQSQTCCQTSQTSIGAF"
complement(8456. .9043)
/gene="tlr11729"
complement(8456. .9043)
/gene="tlr11729"
/note="ORF_ID:tlr11729"
similar to hert protein"
/codon_start=1
/transl_table=1
/protein_id="BAC09281.1"
/db_xref="GI:22295455"
/translation="MMDCPPLTVYQMRAPHRNLTLDSSALILKMLTLPPDVPRQVI
LRAVLRKIQPNLGERPLPRAAGKPYTLNGLFPMWSHSGNLAVLAVSGRAVGVDEIL
RLCPQRAISRRFFGALDQRILEGDSRFLQAMTYEAMLKAGIGVQWRPAAOSLE
HMVASEPVGDRAIASVVLTPPPQCFEFPPEAMA"
complement(9050. .9559)
/gene="tlr11730"
complement(9050. .9559)
/gene="tlr11730"
/note="ORF_ID:tlr11730"
hypothetical protein"
/codon_start=1

FEATURES	Location/Qualifiers
Source	1. 930 /organism="Synecocystis sp. PCC 6803" /db_xref="taxon:1148" 10. 915 /note="unnamed protein product"
CDS	/codon_start=1 /transl_table=11 /protein_id="CAC28417.1" /db_xref="GI:127110085" /translation="MAAGXILFLARQYSSDSVANAAYDQWTEGILEIYWGIGKSPVAKDIPDQSKIDPVHAAQMGGLDLPPEYTVLDVGCIGGSSRIILATGTTISPDQYKRATFELTPDYVAKFAVDAALSPFDSFQVWASVAPPHAKLLRVKREKGIILVADWNRDQVRQVPLNFEKPYMRLLDQMSHPARASLEATQIVGEGVTTADMTVPTLPAMLDPTIMQGIIRPGWQIQYIIRFIRKSVRMLRAFGVGCIRFCFEMFAVRKNAQQA"
BASE COUNT	205 a 230 c 262 g 233 t
ORIGIN	
Alignment Scores:	
Pred. No.:	2.37e-25
Score:	392.00
Percent Similarity:	51.76%
Best Local Similarity:	36.08%
Query Match:	21.42%
DB:	6 Matches: 92 Conservative: 40 Mismatch: 87 Indels: 36 Gaps: 10
US-09-857-613A-28 (1-350) x AX073663 (1-930)	
QY	69 ILEALAGLunPhetYraspgiu-----SerSerglyLeutrpqluanslletrpolyasp 86
Db	64 GTGGCCACAGCCCTACGACCAATGACAGACAGGAGCGGATTTGGATTACTGCGGCGAC 12
QY	87 HisMetHisisclypethYrasSerSerserThnValserLeuSeraspHisargla 10
Db	124 CATATCCAGCTCGGCCATATT---GGCGATCCGCCAGTGCGCCACAGAT----- 16
QY	107 AlagHlIleargMetIleingInleuAurghpeaIasSerValserGlucluarSer 12
Db	169 -----TTCATCCAAATCGAAMAATGATTTTGTG-----CATGCCATGGCC 20
QY	127 LysTrp-----ProlysSerIleValaspValGlycysly 13
Db	208 CAGTGGGGCGGATTGATATCACTTCCCGCCGACACAGGATTTGGATGGGTTGCGCGC 26
QY	139 IleglyGlySerSeraYtyrLeuAlaLysLysPheglYalaThrSeraVIGlylethr 15
Db	268 ATTTGGCGGTAGCAGTCCGATCTCTGCCAAGAATTATGTTTAACTGATCCGGCATCAC 32
QY	159 LeuSerProValGlnAlaGlnArgAlaLysAlaLysAlaAlaAlaGlnGlyLeuAlaasp 17
Db	328 ATTTAGTCCCAACAGTGAACGGGGCGGCGGAAATTAATCTCCCGGATGTGACGGCC--- 38
QY	179 LysValSerPheGlnValAlaAspAlaLeuGlnInPProPheSeraspGlyGlnPheasp 19
Db	385 -----AAGTTTGGCGGTGACGAGATGCTARAGGCTTTGTTCTTCTGACGATGTTTGAC 43
QY	199 LeuValTrpSerMetGluSerGlyLysIleMetProaspLysAlaLysPheValGlyGlu 21
Db	439 GTAGTTTGTGTGGTGAAGCAGGGCGCCACATGCTGCACAAAGCTGTTTGGCCAGGAA 49
QY	219 LeuAlaArgValAlaAlaProGlyAlaIleIleIleValThrTrpCysHisArgasp 23
Db	499 TTACTGCGGGTTCGTAACACAGGGGCGATTCGTGTGTGCGGATTTGATCAACGGGAC 55
QY	239 LeuGlyProaspGluIn-----SerLeuHisPTrpGluInLinsPheLeuLysIle 25
Db	559 -----GATGCCCAAGTGGCCCTCAACTTGTGGAAAAAACAGAGATGCCACAACTG 60
QY	258 CysaspAlaArgtyrtyrLeuProAlaTrpCysSerThrSeraSprtyrValLysLeuGln 27
Db	610 TTGGATCAATAGTCCACACCTGCTTGTGACCATTTGAAGGTTTTCGCGGAATAATTCGAA 66

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 19, 2003, 05:11:11 ; Search time 221 Seconds

(without alignments)
3566.513 Million cell updates/sec

Title: US-09-857-613A-28

Perfect score: 1830

Sequence: 1 MATVVRIPRTSCIHITFRS.....IEGYKDKLIFKAITCRKPE 350

Scoring table:

BIOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODEL=frame_plus_p2n -DB=cgn2_1/USPTO-spool/US09857613/unat_12032003_100035_6003/app_query.fasta_1.519
-DB=N_Geneseq_101002 -QFWT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPECL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BIOSUM62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09857613 -CGN_1_1_200 -runat_12032003_100035_6003 -NCPU=6 -ICPU=3
-NO_XLPRY -NO_MMAR -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV=TIMEOUT=120
-WARR_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_101002: *
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: *
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: *
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: *
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1830	100.0	1189	21	AAA29154	Soybean gamma-toco
2	1668.5	91.2	1331	21	AAA29154	Soybean gamma-toco
3	1173	64.1	1257	21	AAA29165	Soybean gamma-toco
4	1162	63.5	1102	21	AAA29161	Corn gamma-tocophe
5	1154.5	63.1	1293	21	AAC50535	Arabidopsis thalia
6	1151.5	62.9	1071	21	AAZ8317	A. thaliana gene 1
7	1151.5	62.9	1350	24	AAL46039	Arabidopsis gamma-
8	1151.5	62.9	1790	20	AAE17789	Arabidopsis thalia
9	1147.5	62.7	1296	21	AAC34022	Physcomitrella pat
10	1022.5	55.9	1862	22	AAH44261	Wheat gamma-tocoph
11	956	52.2	1011	21	AAA29155	Corn gamma-tocophe
12	778	42.5	792	21	AAA29151	Synechocystis gamm
13	647.5	35.4	954	20	AAE17788	Synechocystis PCC6
14	392	21.4	930	22	AAE26173	DNA encoding a met
15	392	21.4	957	22	AAE26159	Synechocystis PCC6
16	392	21.4	957	22	AAE26169	Synechocystis 2-me
17	392	21.4	957	24	AAE26170	Partial gamma-toco
18	392	21.4	974	22	AAE26170	Corn tassal-derive
19	344.5	18.8	521	21	AAA29152	Complete nucleotid
20	344.5	18.8	521	21	AAA29162	S. lavenulae MltM
21	316	17.3	311	24	ABL74098	Rice gamma-tocophe
22	287	15.7	53500	21	AAC55842	Physcomitrella pat
23	280.5	15.3	852	21	AAC55802	S. longisporiflavus
24	267.5	14.6	464	21	AAA29163	Oxidoreducing aver
25	265	14.5	488	22	AAH44245	A. halophila betai
26	261	14.3	2122	18	AAE70152	Rice gamma-tocophe
27	236	12.9	1410	15	AAO80524	Maize C-24 sterol
28	232.5	12.7	2814	21	AAZ50960	Arabidopsis thalia
29	228.5	12.5	592	21	AAA28153	E. halophilus bet
30	227	12.4	1383	21	AAZ26164	S. spinosa DNA fra
31	226	12.3	1149	21	AAC44838	S. spinosa DNA fra
32	226	12.3	3219	21	AAZ50961	DNA fragment of Sa
33	225	12.3	6085	18	AAE70153	Nicotiana tabacum
34	224.5	12.3	1497	20	AAV70839	Arabidopsis thalia
35	220.5	12.0	828	22	AAE88318	Arabidopsis thalia
36	220.5	12.0	45624	22	AAE88315	Arabidopsis thalia
37	220.5	12.0	50000	22	AAE88312	Arabidopsis thalia
38	220.5	12.0	80161	20	AAZ21501	Arabidopsis thalia
39	218	11.9	12381	21	AAZ58381	Arabidopsis thalia
40	207.5	11.3	1041	24	AAE23672	Arabidopsis thalia
41	204.5	11.2	1233	21	AAC48513	Arabidopsis thalia
42	203.5	11.1	1235	21	AAC39342	Arabidopsis thalia
43	203.5	11.1	1270	21	AAC48512	Arabidopsis thalia
44	203.5	11.1	1274	21	AAC39297	Arabidopsis thalia
45	203.5	11.1	1320	20	AAV70837	Yeast SMT gene. S

ALIGNMENTS

RESULT 1
ID AAA29164 standard; CDNA: 1189 BP.

AC AAA29164;
XX 12-SEP-2000 (first entry)

DE Soybean gamma-tocopherol methyltransferase CDNA (clone sahlc.pK004.g2).

XX Vitamin E; alpha-tocopherol; biosynthesis; enzyme; inhibitor; herbicide;

KW gamma-tocopherol methyltransferase; ss.

XX Glycine max.

OS Key
XX Location/Qualifiers
FH 5..1057
FT CDS
FT /*tag= a

/product= gamma-tocopherol_methyltransferase

FT XX
PN XX WO200032757-A2.
PD XX 08-JUN-2000.
PF XX 02-DEC-1999; 99WO-US28588.
PR XX 03-DEC-1998; 98US-0110781.
XX
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
PI Cahocon RE, Coughlan SJ, Miao G, Rafalski JA.
XX
XX WPI: 2000-412309/35.
DR P-PSDB; AA196472.
XX
XX Polyuclotide encoding a vitamin E biosynthetic enzyme polypeptide
PT such as gamma tocopherol methyltransferase, useful for synthesizing
PT vitamin E or as a target for design and discovery of herbicides

PS Claim 3: Page 64-65; .82pp; English.

XX
XX AAA29151-69 encode vitamin E (alpha-tocopherol) biosynthetic enzymes or
CC their fragments. The cDNA clones were identified by BLAST searches based
CC on similarity to Synecocystis sp. and Arabidopsis thaliana vitamin E
CC biosynthetic enzymes. The enzymes are useful for synthesizing
CC plastoquinones or tocopherols (especially vitamin E). Vitamin E is
CC required for plant growth, therefore the enzymes may be used for the
CC discovery of new herbicides. The enzymes can be used in methods to
CC evaluate potential inhibitors, which may have use as herbicides.
CC Additionally, transgenic expression of, e.g. gamma-tocopherol
CC methyltransferase affords the ability to manipulate tocopherol levels as
CC desired for a particular application.

XX
SQ Sequence 1189 BP; 333 A; 253 C; 299 G; 304 T; 0 other:

Alignment Scores:

Pred. No.:	2,34e-180	Length:	1189
Score:	1830.00	Matches:	350
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	21	Gaps:	0

US-09-857-613a-28 (1-350) x AAA29164 (1-1189)

QY 1 MetalathValAlaArgIleProThrIleSerCysIleHisIleHisThrPheArgSer 20
DB 5 ATGCCACCGTGGTGAAGATCCCAACATCTCATCCATCCACATCCGCTTCC 64
QY 21 GluSerProArgThrPheAlaArgIleArgValGlyProArgSerTrpAlaProIleArg 40
DB 65 CAATCCCTCCGCTTCGCCAGATCCGGTCCGAGCCAGCTCTGCGCTCTATTCCG 124
QY 41 AlaserAlaAlaserserGluArgGlyIleValIleGluGluGlnIleProIleLysAsp 60
DB 125 GCATCCGAGCGAGCTCCGAGAGAGGAGATGATTTGGACCAAGCCGGAAGAGAT 184
QY 61 AspLysLysLysLeuGlnLysGlyIleAlaGluPheTrpAspGluSerSerLysLeuTrp 80
DB 185 GACAAAGAAAGAGCTGCACAAAGGATCCAGAGTTTACAGAGATCCGCTGCTTATCG 244
QY 81 GluAsnIleTrpGlyAspHisMetHisLysGlyPheTrpAspSerAspSerThrValSer 100
DB 245 GAGAACATTTGGGGGACACATGCACATGCTTTTATGATCCGATCCACTCTTCG 304
QY 101 LeuSerAspHisArgAlaAlaGlnIleArgMetIleGlnIleuSerLeuArgPheAlaSer 120
DB 305 CTTCGATCACTCGCTGCTCAGATCCGAATGATCCAGAGCTCTCTTCGCTTGCCTCT 364
QY 121 ValSerGluGluArgSerLysTrpProLysSerIleValAspValGlyCysGlyIleGly 140

DB 365 GTTTCGAGGAGCGTACTAAATGGCCCAAGATATAGTTGATGTGGGTGTGCATAGCT 424
QY 141 GlySerSerArgTrpIleAlaAlaLysLysPheGlyValAlaThrSerValGlyIleHisSer 160
DB 425 GGCAGCTCTGATACCTGGCCAGAGAAATTTGGACCAACCTGTAGGCATCCTAGT 484
QY 161 ProValGlnAlaGlnArgAlaAsnAlaLeuAlaAlaIleGlnIleValAlaAspLysVal 180
DB 485 CTTGTTCAAGCTCAAAAGAGAAATGCTTGCTGCTGCTGCTCAAGATGGCTGATAGCTT 544
QY 181 SerPheGlnValAlaAspAlaLeuGlnIleProPheSerAspGlyGlnPheAspLeuVal 200
DB 545 TCTCTTCAAGTTGCTGAGCGCTCTACACCAACATCTCTGACGCGCATTTGATCTGGTG 604
QY 201 TrpSerMetGluSerGlyGlnHisMetProAspLysAlaLysPheValGlyLeuAla 220
DB 605 TGGTCCATGAGAGATGGAGCATATGCTGACAAAGCTATAGTTTGGAGAGATTAGCT 664
QY 221 ArgValAlaAlaProGlyAlaIleIleIleValAlaThrTrpCysHisArgAspLeuGly 240
DB 665 CGGCTAGAGAGCAGCGTGCAT 724
QY 241 ProAspGluGlnSerLeuHisProTrpGluGlnAspLeuLysLysIleCysAspAla 260
DB 725 CCTGACGACATCTTACATCATGAGGAGCAGATCTTTAAAGAGATTGGAGTCA 784
QY 261 TyrTrpLeuProAlaTrpCysSerThrSerAspTrpValLysLeuGlnSerLeuSer 280
DB 785 TATTACCTCCCTGCTGCTGCTCACTTCTGATTTATGTTAGTTGCTCCATCTCTGCA 844
QY 281 LeuGlnAspIleLysSerGluAspTrpSerArgPheValAlaProPheTrpProAlaVal 300
DB 845 CTTGACGACATCAAGTGAAGATGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 904
QY 301 IleArgSerAlaPheThrTrpLysGlyLeuSerSerLeuLeuSerSerGlyGlnLysThr 320
DB 905 ATACGCTCAGCTTCACATGAGAGGCTATCTTCACTTGCAGCAGCAAAAAACG 964
QY 321 IleLysGlyAlaLeuAlaMetProLeuMetIleGluGlyTrpLysAspLeuLysThr 340
DB 965 ATAAAGAGAGCTTTGGCTATGCTATGCTATGATGATAGAGGATACCAAGAAATCTAATTAG 1024
QY 341 PheAlaIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 1054
DB 1025 TTTGCCATCATTTACATGTCGAAACCTGCA 1054

RESULT 2
AAA29154
ID AAA29154 standard; cDNA: 1331 BP.
AC AAA29154;
XX
XX 12-SEP-2000 (first entry)
DB
DE Soybean gamma tocopherol methyltransferase contig cDNA.
XX
XX Vitamin E: alpha-tocopherol; biosynthesis; enzyme; inhibitor; herbicide;
KW gamma-tocopherol methyltransferase; ss.
XX
XX Glycine max.
OS
XX
XX Key Location/Qualifiers
FT CDS 5..1055
FT /*tag= a
FT /product= gamma_tocopherol_methyltransferase
PN WO200032757-A2.
XX
XX 08-JUN-2000.
PD
XX
XX 02-DEC-1999; 99WO-US28588.
PF
XX
XX 03-DEC-1998; 98US-0110781.

DR P-PSDB: AAY96473.
 XX Polynucleotide encoding a vitamin E biosynthetic enzyme polypeptide
 PT such as gamma tocopherol methyltransferase, useful for synthesizing
 PT vitamin E or as a target for design and discovery of herbicides
 XX
 PS Claim 3, Page 66; 82pp; English.
 XX
 CC AAA29151-69 encode vitamin E (alpha-tocopherol) biosynthetic enzymes or
 CC their fragments. The cDNA clones were identified by BLAST searches based
 CC on similarity to *Synechocystis* sp. and *Arabidopsis thaliana* vitamin E
 CC biosynthetic enzymes. The enzymes are useful for synthesizing
 CC plastoquinones or tocopherols (especially vitamin E). Vitamin E is
 CC required for plant growth, therefore the enzymes may be used for the
 CC discovery of new herbicides. The enzymes can be used in methods to
 CC evaluate potential inhibitors, which may have use as herbicides.
 CC Additionally, transgenic expression of, e.g., gamma-tocopherol
 CC methyltransferase affords the ability to manipulate tocopherol levels as
 CC desired for a particular application.
 CC
 XX
 SQ Sequence 1257 BP; 299 A; 370 C; 343 G; 243 T; 2 other:
 Alignment Scores:
 Pred. No.: 4.76e-112 Length: 1257
 Score: 1173.00 Matches: 224
 Percent Similarity: 80.82% Conservative: 33
 Best Local Similarity: 70.44% Mismatches: 61
 Query Match: 64.10% Indels: 0
 DB: 21 Gaps: 0
 US-09-857-613A-28 (1-350) x AAA29165 (1-1257)
 QY 33 ProAsgSerTrpAlaProLeuArgAlaSerAlaSerSerGluArgGlyLeuVal 52
 DB 117 CCCCCTGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTC 176
 QY 53 LeuGluInLysProLysLysAspLysLysLysLysLysLysLysLysLysLysLys 72
 DB 177 GACGACACCGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 236
 QY 73 TyrAspGluSerSerGlyLeuTrpGluAsnIleTrpGlyAspHisMetHisLysGly 92
 DB 237 TACGACGACTCTCCGCGCTGTGGAGACATCTGGGCGGACACATCTGCACACCGCTTC 296
 QY 93 TyrAspSerAspSerThrValSerLeuSerAspHisArgAlaAlaGlnIleArgMetIle 112
 DB 297 TACGACTCTCGCGGACCGCGCTCCATCTCCGACACCGCGCGCGCGCGCGCGCATGTC 356
 QY 113 GlnGluSerLeuArgPheAlaSerValSerGluLysArgSerLysTrpProLysSerIle 132
 DB 357 GAGGAGCGCCCTCGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTC 416
 QY 133 ValAspValGlyCysGlyIleGlyLysSerSerArgTyrLeuAlaLysLysPheGlyAla 152
 DB 417 GTTGATGTGGATGGGATCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 476
 QY 153 ThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAsnAlaLeuAlaAla 172
 DB 477 CAATGCTCTGCGATCATGATGAGCCAGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 536
 QY 173 AlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAlaAlaLeuGlnIleProPhe 192
 DB 537 GCGCAGGGGTTCGACAGAGCGCTTCTTCAGATGTGTGTGTGTGTGTGTGTGTGTGTGT 596
 QY 193 SerAspGlyLysPheAspLeuValTrpSerMetGluSerGlyGluHisMetProAspLys 212
 DB 597 CCTGATGGCGCAGTTGATCTTGTCTGTATGAGACAGTGGAGACATGCGGACAA 656
 QY 213 AlaLysPheValGlyGluLeuAlaArgValAlaAlaProGlyAlaIleIleIleVal 232
 DB 657 CAGAACTTTGTAAGCAGTGGACCGCTCCGAGCTCCAGGACCACTATCATCATCTG 716
 QY 233 ThrTrpCysHisArgAspLeuGlyProAspGluGlnSerLeuHisProTrpGluGlnAsp 252

DB 717 ACCTGATGCCATAGAGAACTCCGCGCGAGGAGACTCACTGAAACCTGACGCTGAT 776
 QY 253 LeuLeuLysLysIleGlyAspAlaTyrTrpLeuProAlaTrpCysSerThrSerAspTrp 272
 DB 777 CTTTGAAGAAATTTGTATGATCATATTCCTCCGATGTGTGTGTGTGTGTGTGTGTGT 836
 QY 273 ValLysLeuLeuGlnSerLeuSerLeuGlnAspIleLysSerGluAspTrpSerArgPhe 292
 DB 837 GTCAACATTCGCGAGCATGTCTCTTGTAGGATATCAAAACGCCGACGTGCAGAAAC 896
 QY 293 ValAlaProPheTrpProAlaValIleArgSerAlaPheTrpTrpLysGlyLeuSerSer 312
 DB 897 GTGGCCCGCTTCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 956
 QY 313 LeuLeuSerSerGlyLysTrpTrpIleLysGlyAlaLeuAlaIleProLeuMetIleGlu 332
 DB 957 CTACTAAGAGAGTGAAG 1016
 QY 333 GlyTrpLysLysAspLeuIleLysPheAlaIleIleTrpCysArgLysProGlu 350
 DB 1017 GCTTACAGAAAGCGCTCATTAAGTTACAGCATCATCCTGCGGCAAAACCCCA 1070
 RESULT 4
 AAA29161
 ID AAA29161 standard; cDNA; 1102 BP.
 XX
 AC AAA29161:
 XX
 DT 12-SEP-2000 (first entry)
 XX
 DE Corn gamma-tocopherol methyltransferase cDNA.
 XX
 KW Vitamin E; alpha-tocopherol; biosynthesis; enzyme; inhibitor; herbicide;
 KW gamma-tocopherol methyltransferase; ss.
 XX
 OS Zea mays.
 XX
 FT Key Location/Qualifiers
 FT CDS 1..1059
 FT /tag a
 FT /product- gamma-tocopherol_methyltransferase
 FT /transl_except- (pos:454..456, aa:Xaa)
 FT /note- "Xaa not defined"
 XX
 PN MO200032757-A2.
 XX
 PD 08-JUN-2000.
 XX
 PF 02-DEC-1999; 99WO-US28588.
 XX
 PR 03-DEC-1998; 98US-0110781.
 XX
 PA (DUP0) DU PONT DE NEMOURS & CO E I.
 XX
 PI Cahoon RE, Coughlan SJ, Miao G, Rafalski JA;
 XX
 DR WPI: 2000-412309/35.
 DR P-PSDB: AAY96469.
 XX
 PT polynucleotide encoding a vitamin E biosynthetic enzyme polypeptide
 PT such as gamma tocopherol methyltransferase, useful for synthesizing
 PT vitamin E or as a target for design and discovery of herbicides
 XX
 PS Claim 3, Page 60; 82pp; English.
 XX
 CC AAA29151-69 encode vitamin E (alpha-tocopherol) biosynthetic enzymes or
 CC their fragments. The cDNA clones were identified by BLAST searches based
 CC on similarity to *Synechocystis* sp. and *Arabidopsis thaliana* vitamin E
 CC biosynthetic enzymes. The enzymes are useful for synthesizing
 CC plastoquinones or tocopherols (especially vitamin E). Vitamin E is
 CC required for plant growth, therefore the enzymes may be used for the
 CC discovery of new herbicides. The enzymes can be used in methods to

CC evaluate potential inhibitors, which may have use as herbicides.
 CC Additionally, transgenic expression of, e.g. gamma-tocopherol
 CC methyltransferase affords the ability to manipulate tocopherol levels as
 CC desired for a particular application.

XX Sequence 1102 BP; 231 A; 330 C; 328 G; 207 T; 6 other:

Alignment Scores:
 Pred. No.: 5.5e-111 Length: 1102
 Score: 1162.00 Matches: 225
 Percent Similarity: 75.88% Conservative: 33
 Best Local Similarity: 66.18% Mismatches: 72
 Query Match: 63.50% Indels: 10
 DB: 21 Gaps: 2

US-09-857-613A-28 (1-350) x AAA29161 (1-1102)

```

QY 20 SerGlnSerProArgThPheAlaArgIleArgValGlyProArgSerTrpAlaProIle 39
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 28 TCCGAGTCTCCAGAGCCCTCCGAGCTGCGCCGCGCCGAGCTACCGCCCTTCG 87
QY 40 ArgAlaSerAlaAlaSerSerGlnArgGlyIleValLeuGlnGlnLysProLysLys 59
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 88 CACGTCGCCGCCACTCCCGCCGCTCCGAGCGCCGCTGCTAGCCCTGCGCTCGAGGCC 147
QY 60 AspAspLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 71
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 148 TCGTCGACGGCTCAGCGCCCGCCGCGCGCCGCGCTCAGAGAGGCGATCCGCGGG 207
QY 72 PheTrpAspGlnSerSerGlyLeuTrpGlnAsnIleTrpGlyAspHisMetHisGly 91
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 208 CTGTCAGCAGAGTCTCTCGGGCTGTGGAGAACATCTGGCGCGCACATGACACACGCG 267
QY 92 PheTrpAspSerAspSerThrValSerLeuSerAspHisArgAlaAlaGlnIleArgMet 111
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 268 TTCTACGACTCGACGAGCGCCCTCCATGCGCATCGCCGCGCGAGATCCGATG 327
QY 112 IleGlnGlnSerLeuArgPheAlaSerVal-----SerGlnGlnArgSerLysTrpPro 129
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 328 ATCGAGAGGCGCGCTCGCTTCGCGCGCTCCGAGCTCAGATGATCCAGAGAGACACA 387
QY 130 LysSerIleValAspValGlyCysGlyIleGlyGlySerSerArgTrpLeuAlaLysLys 149
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 388 AAACAAATAGTCATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATG 447
QY 150 PheGlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAsnAla 169
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 448 TACGGAGCAGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATG 507
QY 170 LeuAlaAlaAlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAlaLeuGln 189
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 508 CTGCGTCGAGCGCAGCGGTTGTCGATCAGCTTCTGCAAGTTGCTGATCTCTCGAG 567
QY 190 GlnProPheSerAspGlyGlnPheAspLeuValTrpSerMetGlnSerGlyLysHisMet 209
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 568 CAACCGTTTCTGACGCGCAGTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 627
QY 210 ProAspLysAlaLysPheValGlyIleLeuAlaArgValAlaAlaProGlyAlaIleIle 229
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 628 CCGGCAAGAGAAATGTTGTTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTA 687
QY 230 IleIleValThrTrpCysHisArgAspLeuGlyProAspGlnGlnSerLeuHisProTrp 249
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 688 ATCATCGTCAGCATGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 747
QY 250 GlnGlnAspLeuLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 269
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 748 GAATGACGACTCTCTGAGAGATATGACAGCGGTACTACTCTCCGAGCTGGTCTACCT 807
QY 270 SerAspTrpValLysLeuLeuGlnSerLeuSerLeuGlnAspIleLysSerLysTrp 289
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 808 TCAGACTATGTGAACTTCGCAAGTCTCTCTGAGGATATCAAGACAGCAGCTGATCG 867

```

```

QY 290 SerArgPheValAlaProPheTrpProAlaValIleArgSerAlaPheThrTrpLysGly 309
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 868 TCGAGAAACGTGCGCCGCTTTGGCCCGCCGCTGATGATGATGATGATGATGATGATGAT 927
QY 310 LeuSerSerLeuLeuSerSerGlyGlnLysTrpIleLysGlyAlaLeuAlaMetProLeu 329
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 928 TTCACCTCTCTGTCGACGACCGCATGATGATGATGATGATGATGATGATGATGATGAT 987
QY 330 MetIleGlnGlyTrpLysLysAspLeuIleLysPheAlaIleIleThrCysArgLysPro 349
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 988 ATGATTCAGGCGCTACAGAGAGGCGCTCATCAATTCACCATCATCATCTGTCGCAAGCCT 1047
RESULT 5
AAC50535
ID AAC50535 standard; DNA; 1293 BP.
XX
AC AAC50535;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 65205.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
PN BP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 09-MAR-1999; 99US-0123180.
PR 05-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.

```

PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 4,15e-110
Score: 1154.50
Percent Similarity: 75.43%

Length: 1293
Matches: 231
Conservative: 30

Best Local Similarity:	66.76%	Mismatches:	65
Query Match:	63.09%	Indels:	21
DB:	21	Gaps:	4

US-09-857-613A-28 (1-350) x AAC50535 (1-1293)

Oy	8	Proh1IleSerCysIleHisIleHisThrPheArgSerGlnSerProArgThrPheAla	27
Db	118	CGACGCTTCCTTCCGGCTCAAGATCATCGCTTCCTTT-CGGTCCATCTCCTCTCTCC	176
Oy	28	ArgIleArgVal---GlyProArgSerThrPalaProIleArgAlaSerAlaAlaSer	46
Db	177	TCAGTCTCATACGACACACCGCTGGAAACGGCGTGGCGGCTGCTCATACCAT	236
Oy	47	GluArgGlyCuuIleValLeuGlnLysProLysLysAspLysLysLysLeuGln	66
Db	237	GAG-----GCGCTAAGA	248
Oy	67	LysGlyIleAlaGluPheTyrAspGluSerSerGlyLeuTrpGluAsnIleTrpGlyAsp	86
Db	249	AAAGCATACCGAGCTTCACATCAATCAACTCGGTTGTGTGGAAAGATTTGGCGAGAT	308
Oy	87	HisMetHisIleCysPheTyrAspSerAspSerThrValSerLeuSerAsp-----His	104
Db	309	CATATGCATCATCGCTTTTATGACCCCTGATCTTCTCTGTCACATTTCTGATTCGTCAC	368
Oy	105	ArgAlaIleGlnIleArgMetIleGlnGluSerLeuArgPheAlaSerValSer---Glu	123
Db	369	AAGAGACTGAGTCCGATGATGATGAAGGTCTCCGCTTTGGCCGCTTACTGATGAA	428
Oy	124	GluArgSerLysTrpProLysSerIleValAspValGlyCysGlyIleGlyLysSer	143
Db	429	GAGGAGGAGAAAAGATTAAGAAGATGATGTTGGGTGGGATTTGAGAGAACTCA	488
Oy	144	ArgTyrLeuAlaLysLysPheGlyAlaThrSerValGlyIleThrLeuSerProValGln	163
Db	489	AGATATCTTCCCTCAATATTTGGAGCTGAAGCATGGCATTTACTCTCAGCCCTGTTCAG	548
Oy	164	AlaGlnArgAlaAsnAlaLeuAlaAlaAlaGlnGlyLeuAlaAspLysValSerPheGln	183
Db	549	GCCAGAGAGCCATGATCTCGGGCTGCTCAATACCTCGCTCATAGGCTTCTTCCA	608
Oy	184	ValAlaAspAlaLeuGlnGlnProPheSerAspGlyGlnPheAspLeuValTrpSerMet	203
Db	609	GTTCCGATCGTGGATCGCATCGAATGAGAAATTCGATCTAGTGTGTCGATG	668
Oy	204	GluSerLysGlyLysMetProAspLysAlaLysPheValGlyGluLeuAlaArgValAla	223
Db	669	GAGAGTGTGAGCATATGCTCTGCAAGGCCAAGTTGTGAAAAAGATTTGTCAGTGTGGCG	728
Oy	224	AlaProGlyAlaIleIleIleIleValIleThrTrpCysHisArgAspLeuGlyProAspGlu	243
Db	729	GCTTCAGAGAGTAGATATAATATAGTACATGTGTCCATAGAAATCTATCTGCGGGAG	788
Oy	244	GlnSerLeuHisProTrpGlnGlnAspLeuLysLysIleCysAspAlaTyrTrpLeu	263
Db	789	GAACCTTTGACGCGGTGGAGCAAAACATCTTGAGCAAAATCTGTAAAGACCTCTATCTC	848
Oy	264	ProAlaTrpCysSerThrSerAspTyrValLysLeuLeuGlnSerLeuSerLeuGlnAsp	283
Db	849	CCGGCTGTGGCTCCACCGATGTATTTGTCACTGTGCTTCATATCCCATTTCTCTCCAGAT	908
Oy	284	IleLysSerGluAspTrpSerArgPheValAlaProPheTrpProAlaValIleArgSer	303
Db	909	ATTAAAGTGTGGATTTGTCAGAACCTGACCTCTTCTGTGGCCGCGGTATACGAGCT	968
Oy	304	AlaPheThrTrpLysGlyLeuSerSerLeuLeuSerSerGlyGlnLysThrIleLysGly	323
Db	969	GCATTACATGAGAGGCGCTTGTCTGCTCTGCTTCGATGAGTATGAAGATTTAAAGCA	1028
Oy	324	AlaLeuAlaMetProLeuMetIleGlnGlyTyrLysLysAspLeuIleLysPheAlaIle	343
Db	1029	GCATTGACAAATGCCATTCGATGATTTGAAGGTTCATAAAGAGGTGCTCAATTAAGTTTGATATC	1088

QY	344	IIEThrcysargLysPro	349
		:::	
Db	1089	ATCACTGCCAGAAGCCA	1106

RESULT 6

ID AAZ98317 standard; DNA; 1071 BP.

AC AAZ98317

DT 14-JUN-2000 (first entry)

A. thaliana gene involved in environmental stress tolerance.

KW Environmental stress; plant; transgenic plant; anaerobic; flooding; cold
KW dehydration; drought; heat stress; salinity; osmotolerance; ds.

OS *Arabidopsis thaliana*.

PN WO200008187-A2

PD 17-FEB-2000.

PF 04-AUG-1999; 99WO-EP05652.

PR 04-AUG-1998; 98EP-0202634.

PA (VLA-) VLAMS INTERUNIVERSITAIR INST BIOTECHNOG.
VU

PI Lee JH, Verbruggen N;
xx

DR WPI; 2000-205726/18.

[illegible]

PT isolating genes involved in tolerance to environmental stress

PS Claim 4; Page 130-132; 312pp; English.

CC The inversion relates to isolation of coding sequences and/or genes
CC involved in tolerance to environmental stress in plants. The sequences
CC (AAZ98305-298365) are useful for producing a transgenic plant having
CC enhanced tolerance or resistance to environmental stress conditions such
CC as anaerobic, flooding, cold, dehydration, drought, heat stress or
CC salinity. This is useful for producing improved yield, growth,
CC development and productivity under environmental stress conditions, and
CC also provides growth of crops in areas where they cannot grow without
CC the induced osmotolerance. Sequences AAZ98305-365 represent
CC polynucleotide sequences from *A. thaliana* that are involved in
CC environmental stress tolerance.

SQ Sequence 1071 BP; 277 A; 229 C; 270 G; 295 T; 0 other;

Alignment Scores:

Score:	1151.50	Matches:	230
--------	---------	----------	-----

Best Local Similarity: 66.47% Mismatches: 6

DB:	21	Gaps:	4
-----	----	-------	---

US-09-857-613A-28 (1-350) x AA298317 (1-1071)

8 ProthrIleSerCysIleHisIleThrPheArgSerGlnSer

Db 61 CCACTCTTCTTCCGGCTCAAGTCATCGCTTCTC

```

QY 28 ArgIleArgVal--GlyProArgSerTrpAlaProIleAlaGlaSerAlaSer 46
      ::::: ||| ::::: |||::: |||:::
Db 120 TGAAGTCCTATATACACACACGGGTGGAAACGGGTGGCGCGCTGCTACATCCACT 170
      ::::: ||| ::::: |||::: |||:::
QY 47 GluArgGlyGluIleValLeuCluGlnLysProLysLysAspAspLysLysLysLeu 66

```

```

180 GAC-----GGCTAAGA 191
OY 67 LysGlyLeuAlaGlnPheTyrAspGluSerSerGlyLeuTrpGluAsnIleTrrpGlyAsp 86
DB 192 AAGGATATGCGGAGTTCTCAATGAAGAACTTCGGGTTTGCGGAGAGATTGGGAGAT 251
OY 87 HisMetHisIleGlyPheTyrAspSerSerThrValSerLeuSerAsp-----His 104
DB 252 CATATCATCATGATGCTTTATATGACCGCTGATCTCTGTTCACTTCTGATTTCTGCTCAC 311
OY 105 ArgAlaAlaGlnIleArgMetIleGlnGluSerLeuArgPheAlaSerValSer---Glu 123
DB 312 AAGGAGCTCAGATCCGTTATGATGAAGAGTCTCCGTTCCCGGCTTACTGATGA 371
OY 124 GluArgSerLysTrpProLysSerIleValAspValGlyCysGlyIleGlyGlySerSer 143
DB 372 GAGGAGCAGAAAGATAAAGAAAGTAGATGTTGGTGCGTGGATGGAGAGAACTCA 431
OY 144 ArgTyrLeuAlaLysLysPheGlyAlaThrSerValGlyIleThrLeuSerProValGln 163
DB 432 AGATATCTTGCTCTTAATTTGGAGCTGAATGATGCTTCTCTCAGCCCTGTTCA 491
OY 164 AlaGlnArgAlaAsnAlaLeuAlaAlaGlnGlyLeuAlaLysPheValSerPheGln 183
DB 492 GCCAAGAGAGCCAAATGATCTCGCGCTGCTCATATCACTCTCATAGGCTTCTTCAA 551
OY 184 ValAlaAspAlaLeuGlnTrpPheSerAspGlyGlnPheAspLeuValTrrpSerMet 203
DB 552 GTTGGCGGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 611
OY 204 GluSerGlyLysIleHisMetProAspLysAlaLysPheValGlyLeuAlaArgValAla 223
DB 612 GAGAGTGGAGATATGCTGCTGACAGCCAACTTCTTAAGAGCTTGGTACGTGGCG 671
OY 224 AlaProGlyAlaIleIleIleIleValIleThrPheCysHisArgAspLeuGlyProAspGlu 243
DB 672 GCTCCAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 731
OY 244 GlnSerLeuHisProTrpGluAsnAspLeuLysIleCysAspAlaTyrTrrpLeu 263
DB 732 GAAGCTTTCAGCCGCTGGAGACCAAAACATCTTGACAAATCTGTAAAGCTTCTATCTC 791
OY 264 ProAlaTrpCysSerThrSerAspTyrValLysLeuLeuGlnSerLeuSerLeuGlnAsp 283
DB 792 CCGGCTTGGGCTCCACCGATGATGATGATGATGATGATGATGATGATGATGAT 851
OY 284 IleLysSerGluAspTrpSerArgPheValAlaProPheTrpProAlaValIleArgSer 303
DB 852 ATTAAGTGTGCGGATGCTGAGAGAACGTAAGCTCTTCTGCGCTGCGGTTATACGACT 911
OY 304 AlaPheThrTrpLysGlyLeuSerSerLeuLeuSerSerGlyLysIleThrIleLysGly 323
DB 912 GCATTTACATGGAAGGCGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 971
OY 324 AlaLeuAlaMetProLeuMetIleGlnGlyTyrLysAspLeuIleLysPheAlaIle 343
DB 972 GCATTACATGCGCATTTGATGATGAAGTTACAGAAAGCTGTCTATTAGTTGGTATC 1031
OY 344 IleThrCysArgLysPro 349
DB 1032 ATCACTTCCAGAACCA 1049

```

RESULT 7
AAL46039
ID AAL46039 standard: cDNA: 1350 BP.

AC AAL46039;
XX 11-JUL-2002 (first entry)
DT
DE A thaliana gamma-tocopherol methyltransferase cDNA.
XX

```

KW Vitamin E; homogenitase; HG: homogenitase-1,2-dioxygenase; HGD:
KW malylacetate-1-isomerase; MAAI: fumarylacetate-1-hydrolyase; FAAH:
KW transgenic; nutrition; antioxidant; cardiovascular disease; cancer;
KW immune system; generalised age-related degeneration; animal feed;
KW meat quality; cosmetics; growth regulator; herbicide; cardiant;
KW cytosolic; immunostimulant; enzyme; gene; ss.
OS Arabidopsis thaliana.
FH key Location/Qualifiers
FT CDS 63..1109
FT /tag= a
FT /product= "gamma-tocopherol methyltransferase"
PN MO200231173-A2.
PD 18-APR-2002.
PF 18-SEP-2001; 2001WO-EP10779.
PR 19-SEP-2000; 2000DE-1046462.
PA (SUNG-) SUNGENE GMBH & CO KGAA.
PI Geiger M, Ebneith M, Kunze I.
PI WPI: 2002-362682/39.
DR P-PSDB: AAO17426.
XX
XX Increasing Vitamin E production, useful in human or animal nutrition,
XX comprising reducing activity of enzymes that metabolize homogenitase
XX Disclosure: Page 83-85; 98pp; German.
XX
XX The present invention relates to a method of producing Vitamin E,
XX involving modulating the synthesis of vitamin E by reducing decomposition
XX of homogenitase (HG) by reducing the activity of
XX homogenitase-1,2-dioxygenase (HGD), malylacetate-1-isomerase (MAAI)
XX and/or fumarylacetate-1-hydrolyase (FAAH). The method is used to
XX produce transgenic organisms that are useful in animal and human
XX nutrition and for the isolation of vitamin E. Vitamin E is a fat-soluble
XX antioxidant with a protective effect against cardiovascular disease and
XX cancer, it also stimulates the immune system and may prevent generalised
XX age-related degeneration. When used in animal feeds, it improves quality
XX and storability of meat and can also be used in cosmetics. Also
XX antibodies raised against the 3 specified enzymes are useful in screening
XX for specific inhibitors, potentially useful as growth regulators, e.g.
XX herbicides. The present sequence is the A thaliana gamma-tocopherol
XX methyltransferase cDNA.
SQ Sequence 1350 BP; 374 A; 278 C; 314 G; 384 T; 0 other;

```

Alignment Scores:
Pred. No.: 9,04e-110 Length: 1350
Score: 1151.50 Matches: 230
Percent Similarity: 75.43% Conservative: 31
Best Local Similarity: 66.47% Mismatches: 65
Query Match: 62.92% Indels: 21
DB: 24 Gaps: 4

US-09-857-613A-28 (1-350) x AAL46039 (1-1350)

```

OY 8 ProThiIleSerCysIleHisIleHisThrPheArgSerGlnSerProArgThrPheAla 27
DB 115 CCACACTCTTCTTCCGCTCAAGTCAATCCCTCTCTT-CGGCTTCACACCTCTCTCTC 173
OY 28 ArgIleArgVal---GlyProArgSerTrpAlaProIleArgAlaSerValSerSer 46
DB 174 TCAGTCTCTATACGACGACGCGTGAAGACGTCGCTGCGGCTCTCTCTCATCTCACT 233
OY 47 GluArgGlyLysIleValLeuGlnGlnLysProLysLysAspAspLysLysLysLeuGln 66
DB 234 GAG-----GCCTAAGA 245

```

```

OY 67 LysGlyIleAlaGluPheTyrAspGluSerSerGlyLeuTrpGluAsnIleTrpGlyAsp 86
DB 246 AAGAGATACGCGAGTCTCAATGAAGAACTCGGCTTGTGGAGAGATTGGGAGAT 305
OY 87 HisMetHisIstGlyPheTyrAspSerAspSerThrValSerLeuSerAsp-----His 104
DB 306 CATATGATCATGGCTTTTATGACCCCTGATCTTCACTTCACTTCACTTCACTTCACT 365
OY 105 ArgAlaIleAlaGlnIleArgMetIleGlnIleSerLeuArgPheAlaSerValSer---Glu 123
DB 366 AAGAGACTGATCCGATGATGATGAGAGCTCTCGCTTGGCGGCTGTACTGATGAA 425
OY 124 GluArgSerIleTrpProLysSerIleValAlaPheValGlyGlyIleGlySerSer 143
DB 426 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 485
OY 144 ArgTyrLeuAlaLysIlePheGlyAlaThrSerValGlyIleThrLeuSerProValGln 163
DB 486 AGATATCTTGGCTTAAATTTGAGCAATGCAATGCAATGCAATGCAATGCAATGCA 545
OY 164 AlaGlnArgAlaAsnAlaLeuAlaAlaAlaGlnIleGluLeuAlaAspLysValSerPheGln 183
DB 546 GCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 605
OY 184 ValAlaAspAlaLeuGlnIleGlnIleProPheSerAspGlyIlePheAspLeuValTrpSerMet 203
DB 606 GTTGGCGATCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 665
OY 204 GluSerGlyIleHisMetProAspLysAlaLysPheValGlyIleLeuAlaArgValAla 223
DB 666 GAGAGTGTGAGATATGATGATGATGATGATGATGATGATGATGATGATGATG 725
OY 224 AlaProGlyValIleIleIleIleValThrTrpCysHisArgAspLeuGlyProAspGlu 243
DB 726 GCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 785
OY 244 GlnSerLeuHisProTrpGlnIleAspLeuLysIleCysAspAlaTyrTrpLeu 263
DB 786 GAGCTTTGAGCGCTGGAGCAAAACATCTTGCAAAATCTGTAGAGCTTCTATCTC 845
OY 264 ProAlaTrpCysSerThrSerAspTyrValLysLeuGlnSerLeuSerLeuGlnAsp 283
DB 846 CCGGCTTGTGCTCCACCGATGATGATGATGATGATGATGATGATGATGATGAT 905
OY 284 IleLysSerGluAspTrpSerArgPheValAlaProPheTrpProAlaValIleArgSer 303
DB 906 ATTAAGTGTGCGATGATGATGATGATGATGATGATGATGATGATGATGATG 965
OY 304 AlaPheThrTrpLysGlyLeuSerSerLeuLeuSerSerGlyGlnLysThrIleLysGly 323
DB 966 GCATTTAAGATGAGAGAGCGCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1025
OY 324 AlaLeuAlaMetProLeuMetIleGluGlyTyrLysLysAspLeuIleLysPheAlaIle 343
DB 1026 GCATTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1085
OY 344 IleThrCysArgLysPro 349
DB 1086 ATCACTTGTCCAGAACCA 1103

```

```

XX Arabidopsis sp.
OS
XX Key Location/Qualifiers
FH CDS 207..1253
FT /tag= a
FT /gene= "165H5T7"
FT /product= "gamma-TMT"
XX
XX W09904622-A1.
XX
XX 04-FEB-1999.
XX
XX 22-JUL-1998; 98MO-US15137.
XX
XX 17-JUL-1998; 98US-0053819.
XX 25-JUL-1997; 97US-0053819.
XX 26-JAN-1998; 98US-0072497.
XX
XX (UYNE-) UNIV NEVADA.
XX
XX Delapenna D, Shintani DK;
XX
XX WPI: 1999-142458/12.
XX
XX P-PSDB: AAW95017.
XX
XX Newly isolated DNA fragment comprising a ^c-tocopherol (vitamin E)
XX methyltransferase coding sequence useful for producing
XX ^a-tocopherol, and transgenic plants, seeds and oils with an altered
XX tocopherol profile
XX
XX Claim 2: Page 36-38; 46pp: English.
XX
XX The invention provides DNA sequences encoding gamma-tocopherol
XX methyltransferases (gamma-TMT). The gamma-TMT encoding genes SLR0089
XX and 165H5T7 are isolated from Synechocystis and Arabidopsis species
XX respectively. The DNA fragments are useful for producing transgenic
XX plants containing a higher alpha-tocopherol:gamma-tocopherol ratio. They
XX are also useful for producing alpha-tocopherol and plants with increased
XX gamma-tocopherol, which may be useful in certain industries such as the
XX meat industry e.g. for developing forage plants to feed animals. The
XX production of transgenic plants (and seeds) with a higher alpha-
XX tocopherol:gamma-tocopherol ratio, using the gamma-TMTs, will increase
XX the level of alpha-tocopherol in the human diet, in addition to enhancing
XX the stability and shelf life of plants and plant products. Increased
XX levels of alpha-tocopherol will also increase meat quality and extend
XX shelf life of post-processed meat products. Plants with a higher alpha-
XX tocopherol:gamma-tocopherol ratio may also produce advantageous
XX phenotypes. The present sequence represents the 165H5T7 gene encoding the
XX Arabidopsis gamma-TMT.
XX
XX Sequence 1790 BP; 474 A; 404 C; 436 G; 476 T; 0 other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 1,35e-109 Length: 1790
XX Score: 1151.50 Matches: 230
XX Percent Similarity: 75.43% Conservative: 31
XX Best Local Similarity: 66.47% Mismatches: 65
XX Query Match: 62.92% Indels: 21
XX DB: 20 Gaps: 4
XX
XX US-09-857-613a-28 (1-350) x AAX17789 (1-1790)
XX
XX OY 8 ProThrIleSerCysIleHisIleHisThrPheArgSerGlnSerProArgThrPheAla 27
XX DB 259 CCAACTCTTCTTTCGCTCAAGCTGATCGCTCTT-CCGCTCCATCTCTCTCC 317
XX
XX OY 28 ArgIleArgVal---GlyProArgSerTrpAlaProIleArgAlaSerAlaSerSer 46
XX DB 318 TCACTCTCTATGACGACACACGCGTGAACGCGCTGTGCGCTGCTCATCCTCCT 377
XX OY 47 GluArgGlyGluIleValLeuGlnIleLysProLysLysAspLysLysLysLeuGln 66
XX DB 111

```

RESULT 8
AAX17789
ID AAX17789 standard; DNA; 1790 BP.
AC AAX17789;
XX
XX
XX 21-MAY-1999 (first entry)
DE Arabidopsis gamma-tocopherol methyltransferase encoding gene.
XX
XX Gamma-tocopherol methyltransferase; gamma-TMT; SLR0089; 165H5T7; meat;
KW transgenic plant; alpha-tocopherol; gamma-tocopherol; forage plant; ds.


```

PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143524.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.

```

```

PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 28-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Alignment Scores:

```

Pred. No.: 2.22e-109
Score: 1147.50
Percent Similarity: 75.438
Best Local Similarity: 66.188
Query Match: 62.708
DB: 21
Gaps: 4

```

US-09-857-613A-28 (1-350) x AAC34022 (1-1296)

```

QY 8 ProthrlleSerCysIleHisIleHsrPheArgSerGlnSerProArgThrPheAla 27
DB 120 CCACCTCTTCTTCGCGTCAAGTCATCGCTTCTTTCGCTTCACACCTTCCTCC 178
QY 28 ArgIleArgVal--GlyProArgSerThrAlaPoiIleArgAlaSerAlaIaSerSer 46
DB 179 TCAGTCTATGACGACACACGCGTGAACGTCGCGCGCGCTGCTACATCCACT 238
QY 47 GluArgGlyIuIleValLeuGlnGlnIlyProIlySAspAspLysLysLeuGln 66
DB 239 GAG-----GCCCTAACA 250
QY 67 LysGlyIleAlaGluPheTyraSPgluSerSerGlyLeuTrpGluAsnIleTrpGlyAsp 86
DB 251 AAAGGAATACGCGAGCTTCTACATGAACATTCGCGTTGTGGAGAGATTGGGAGAT 310
QY 87 HlsmelHlshlGlyPheTyraSPSerSerThrValSerLeuSerAsp-----HlS 104

```



```

OY 119 AlaservAlsergluAluagSerLysTrpProLysSerIleValAspValGlyCysGly 138
DB 559 GCTGGCGCTTCTCTACCAAGATTGTAACCGAAGACGATCGTCGATGCGTTGGG 618
OY 139 ILGLYGLYserSerArgTrpLeuAlaLeuLysPheGlyAlaThrSerValGlyIleThr 158
DB 619 ATAGGGGGAAGCTTCACGTTACTTGGCCGGAATTCACGCCCAAGGTAATGCCATCAG 678
OY 159 LeuSerProValGlnAlaGlnAlaAsnAlaLeuAlaLeuAlaGlnGlyLeuAlaAsp 178
DB 679 CTCAGGCCAGTGAGGTTGACAGACGCCGTAAGACCTTACTGACCAAGCAAGCTTATCTGAC 728
OY 179 LysValSerPheGlnValAlaAspAlaLeuGlnGlnProPheSerAspGlyGlnPheAsp 198
DB 739 CTCGTCGAATTTCCAGTACGGAATGCGCTCAACACACCCCTTACAGATGCTGTTGAT 798
OY 199 LeuValTrpSerMetGluSerGlyGlnHisMetProAspLysAlaLysPheValGlyGln 218
DB 799 CTCGCTGCTCCATGAGAGCGCGGACATGCCAGCAAGAAAGTTTGGCGCAG 858
OY 219 LeuAlaArgValAlaAlaProGlyAlaIleIleIleIleValThrTrpCysHisArgAsp 238
DB 859 CTTCGACGAGTAGACGCTCCGCGCGCTGCATTCCTGTCAGCGTGCACCGCTGAT 918
OY 239 LeuGlyProAspGlyGlnSerLeuHisProTrpGlnGlnAspLeuLysIleCys 258
DB 919 CTCAAGCCCGGTAACCTTCTCTCAAGCCTGACGACGATCTTTTGACAAAGATTGT 978
OY 259 AspAlaTrpTrpLeuProAlaTrpCysSerThrSerAspTrpValLysLeuLeuIns 278
DB 979 GACCATTTCTACTTCCGACGCTGCTGCGCGCGGACGCTCATGCTCATGACCAAG 1038
OY 279 LeuSerLeuGlnAspIleLysSerGlnAspTrpSerArgPheValAlaProPheTrp 298
DB 1039 TTAGCGCTGCGAGATCAACAGGAGGCTGTGCTCGAGTACGTCACGCCCTTCTGCGCA 1098
OY 299 AlaValIleArgSerAlaPheThrTrpLysGlyLeuSerSerLeuLeuSerSerGlyGln 318
DB 1099 GCGGTGATGAAACCCGCTTCTCTCATGAAAGGCTGTGGGACGTCGTCAGCTCGGCTCG 1158
OY 319 LysThrIleLysGlyAlaLeuAlaMetProLeuMetIleGlyTrpLysLysAspLeu 338
DB 1159 ACACATCTGAAGAGAGCTTTCGCGATGACCGCTCATGATCAGGCTACAGCGAGGCTC 1218
OY 339 IleLysPheAlaIleIleThrCysArgLys 348
DB 1219 ATTAATTCGCTGCATCACTTCAGAGAG 1248

RESULT 11
AAA29155
ID AAA29155 standard; cDNA; 1011 BP.
XX
AC AAA29155;
XX
DT 12-SEP-2000 (first entry)
XX
DE Wheat gamma tocopherol methyltransferase cDNA.
XX
KW Vitamin E: alpha-tocopherol; biosynthesis; enzyme; inhibitor; herbicide;
KW gamma-tocopherol methyltransferase; ss.
XX
OS Trilicium aestivum.
XX
Key Location/Qualifiers
CDS 3..885
FT /tag= a
FT /product= gamma tocopherol methyltransferase
FT /transl_except= (pos:381..383, aa:Xaa)
FT /transl_except= (pos:387..389, aa:Xaa)
FT /transl_except= (pos:390..392, aa:Xaa)
FT /transl_except= (pos:393..395, aa:Xaa)

```

```

FT /transl_except= (pos:399..401, aa:Xaa)
FT /note= "Xaa is not defined"
XX
XX WO200032757-A2.
XX
XX 08-JUN-2000.
XX
XX 02-DEC-1999; 99WO-US28588.
XX
XX 03-DEC-1998; 98US-0110781.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E. I.
XX
XX Cahoon RE, Coughlan SJ, Miao G, Rafalski JA;
XX
XX MPI: 2000-412309/35.
XX
XX P-PSDB: AAT96463.
XX
XX Polynucleotide encoding a vitamin E biosynthetic enzyme polypeptide
XX such as gamma tocopherol methyltransferase, useful for synthesizing
XX vitamin E or as a target for design and discovery of herbicides
XX
XX Claim 3; Page 53; 82pp; English.
XX
XX AAA29151-69 encode vitamin E (alpha-tocopherol) biosynthetic enzymes or
XX their fragments. The cDNA clones were identified by BLAST searches based
XX on similarity to Synchocystis sp. and Arabidopsis thaliana vitamin E
XX biosynthetic enzymes. The enzymes are useful for synthesizing
XX plastoquinones or tocopherols (especially vitamin E). Vitamin E is
XX required for plant growth, therefore the enzymes may be used for the
XX discovery of new herbicides. The enzymes can be used in methods to
XX evaluate potential inhibitors, which may have use as herbicides.
XX Additionally, transgenic expression of, e.g. gamma-tocopherol
XX methyltransferase affords the ability to manipulate tocopherol levels as
XX desired for a particular application.
XX
XX Sequence 1011 BP; 255 A; 269 C; 268 G; 216 T; 3 other;
XX
XX Alignment Scores:
XX Pred. No.: 1.24e-89 Length: 1011
XX Score: 956.00 Matches: 194
XX Percent Similarity: 79.21% Conservative: 27
XX Best Local Similarity: 69.53% Mismatches: 55
XX Query Match: 52.24% Indels: 4
XX DB: 21 Gaps: 0
XX
XX US-09-857-613A-28 (1-350) x AAA29155 (1-1011)
XX
XX OY 75 GluSerSerGlyLeuTrpGluAsnIleTrpGlyAspHisMetHisGlyPheTrpAsp 94
XX DB 3 GACTCGTCCGCGCTGTGGAGAGCATCTGGGGAGACACATGACACACGCGCTTACGAC 62
XX OY 95 SerAspSerThrValSerLeuSerAspHisArgAlaAlaGlnIleArgMetIleGlnGlu 114
XX DB 63 TCCGGGAGAGCCGCCCTCCATGTCGACACACCGCGCGCCAGATCCGATCGAGAG 122
XX OY 115 SerLeuArgPheAlaSerValSerGlnGlnArgSerLysTrpProLysSerIleValAsp 134
XX DB 123 GCCCTGCGCTTCCGCGCCCTCCCGCAGATCGACAAACAAACCCAAACGATTGTGAT 182
XX OY 135 ValGlyCysGlyIleGlyGlySerSerArgTrpLeu-AlaLysLysPheGlyAlaThrSe 154
XX DB 183 GTTGATGCGGAGATGGGTGAGTCAAGATACCTGGCGGACAAATATGAGGACAAATG 242
XX OY 154 rValGlyIleThrLeuSerProValGlnAlaGlnAlaAsnAlaLeuAlaAlaGln 174
XX DB 243 CTCGCGATCACTTACAC-CCAGTCAAGCTGAGAGGAAATGCGCGCGCAGCGCA 301
XX OY 174 nGlyLeuAlaAspLysValSerPheGlnValAlaAspAlaLeuGlnGlnProPheSerAs 194
XX DB 302 AGGGGTTTCCGGAGCAAGGTTCTTCAATTGCGATCTCGGAGCAACATTTCCTGG 361
XX OY 194 pGlyGlnPheAspLeuVal-TrpSerMetGluSerGlyGlnHisMetProAspLysAla 214

```

```

Db 362 ATGGGATTTGATCTTGTCTGGGNCATGGAGAGANTGTGACACATCGGAGAACACAGA 421
Oy 214 yshvvalglglulealaarvalalalalproglalalellellellevalthrt 234
Db 422 AGTTTGAACGACCTGCGACGCTCCAGAGACTATCATCATGCTGACCT 481
Oy 234 rpyshisargsppluglulproasppluglulserleuhsiprotrpglulnaspplu 254
Db 482 GGTCCATAGAACCTCGCCATCGGAGACTCAGTCAACCGACGAGCTGATCTT 541
Oy 254 eulyslyslelecyasapalalyrtrleuproalalyrtrleuproalalyrtrleu 274
Db 542 TGAATAAATTTGATGATCATATTCACCTCCGAGTGTGCTCTCCGATTAATGCA 601
Oy 274 yslleuenglulserleuhsiprotrpglulnasppluglulserleuhsiprotr 294
Db 602 AGATTGCCAGTCAATCTGTCTTGAAGATATCAAAAGCCGCTGCTGTAAGACCTG 661
Oy 294 laphrohetrproalalyrtrleuproalalyrtrleuproalalyrtrleu 314
Db 662 CCCGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 721
Oy 314 euserserglylulnalyrtrleuhsiprotrpglulnasppluglulserleu 334
Db 722 TAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 781
Oy 334 ylyslasplleuhsiprotrpglulnasppluglulserleuhsiprotrpglul 350
Db 782 ACAAGAAAGGCTCATTAAGTTCAAGCATCATCACTCCACAAACCCCA 832

RESULT 12
AAA29151
ID AAA29151 standard; CDNA: 792 BP.
AC AAA29151:
XX 12-SEP-2000 (first entry)
DE Corn gamma tocopherol methyltransferase partial coding sequence.
XX Vitamin E: alpha-tocopherol: biosynthesis; enzyme; inhibitor; herbicide;
KM gamma-tocopherol methyltransferase; ss.
XX Zea mays.
OS location/Qualifiers
FH Key 2..574
FT CDS /product= gamma_tocopherol_methyltransferase
FT /partial
XX WO200032757-A2.
XX 08-JUN-2000.
XX 02-DEC-1999; 99WO-US28588.
XX 03-DEC-1998; 98US-0110781.
XX (DUPO) DU PONT DE NEMOURS & CO E. I.
XX PA Cahoon RE, Coughlan SJ, Miao G, Rafalski JA:
XX PI WPI: 2000-412309/35.
XX DR P-BSDB: AAY96459.
XX PT polynucleotide encoding a vitamin E biosynthetic enzyme polypeptide
XX PT such as gamma tocopherol methyltransferase, useful for synthesizing
XX PT vitamin E or as a target for design and discovery of herbicides
XX PS Claim 3; Page 49; 82pp: English.
XX

```

```

CC AAA29151-69 encode vitamin E (alpha-tocopherol) biosynthetic enzymes or
CC their fragments. The cDNA clones were identified by BLAST searches based
CC on similarity to Synecocystis sp. and Arabidopsis thaliana vitamin E
CC biosynthetic enzymes. The enzymes are useful for synthesizing
CC plastoquinones or tocopherols (especially vitamin E). Vitamin E is
CC required for plant growth, therefore the enzymes can be used in methods to
CC discovery of new herbicides. The enzymes may be used as herbicides.
CC evaluate potential inhibitors, which may have use as herbicides.
CC Additionally, transgene expression of, e.g., gamma-tocopherol
CC methyltransferase affords the ability to manipulate tocopherol levels as
CC desired for a particular application.
CC
XX Sequence 792 BP; 201 A; 204 C; 215 G; 166 T; 6 other:
SO
Alignment Scores:
Pred. No.: 2,77e-71 Length: 792
Score: 778.00 Matches: 145
Percent Similarity: 85.86% Conservative: 19
Best Local Similarity: 75.92% Mismatches: 27
Query Match: 42.51% Indels: 0
DB: Gaps: 0
US-09-857-613a-28 (1-350) x AAA29151 (1-792)
Oy 159 LeuserProvalglulnalyrtrleuhsiprotrpglulnasppluglulserleu 178
Db 2 TTAGAGCCCTGTTCAAGCCGAGAGAGAAATGCTCTCCCTCAGCGGAGGCTGCGAT 61
Oy 179 LysValSerPheGlnValAlaAspAlaLeuGlnInProPheSerArgPheValGly 198
Db 62 CAGGTACTCTGCAAGTGTGCTGATGCTGCTGAGCAACCGTTCTGAGGAGGCTTCAT 121
Oy 199 LeuValTrpSerMetGlnSerGlyGlnIleMetProAspLysAlaLysPheValGly 218
Db 122 CTGGTGTGCTCATGAGAGGTGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 181
Oy 219 LeuAlaArgValAlaAlaProGlyAlaIlellellellellellellellellelle 238
Db 182 CTAGCAGCGGTGCGGCTCTCTGAGGAGCAATATCATGTCATGATGATGATGATGAT 241
Oy 239 LeuGlyProAspGlnInserleuhsiprotrpglulnasppluglulserleuhsip 258
Db 242 CTGATTCATTCAGAACCTGCTTAACCCGAGTGAAGTGAAGTGAAGTGAAGTGAAG 301
Oy 259 AspaLysTrpLeuProAlaTrpCysSerThrSerArgPheValLysLeuGlnSer 278
Db 302 GACGCTACTACTCTCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
Oy 279 LeuserleuGlnAspLysSerGlnAspTrpSerArgPheValAlaProPheTrp 298
Db 362 CTGCTCTCGAGATATCAAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
Oy 299 AlaValIleArgSerAlaPheTrpTrpLysGlyLeuSerleuhsiprotrpgluln 318
Db 422 GCCGTGATTAATTCAGCCGCTAATGAGAGGCTTCACCTCTCTGCTGCTGCTGCT 481
Oy 319 LysTrpIleGlyAlaLeuAlaMetProLeuMetIleGlyLysLysLysLysLys 338
Db 482 AAGACGATCAGAGCGCGGATGATGATGATGATGATGATGATGATGATGATGAT 541
Oy 339 IleLysPheAlaIlellellellellellellellellellellellellelle 350
Db 542 ATCAATTCACATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 574

RESULT 13
AAAX17788
ID AAAX17788 standard; DNA: 954 BP.
AC AAAX17788:
XX 21-MAY-1999 (first entry)
XX Synechocystis gamma-tocopherol methyltransferase encoding gene.
XX

```

XX Gamma-tocopherol methyltransferase; gamma-TMT; SLR0089; 165H5T7; meat;
 KW transgenic plant; alpha-tocopherol; gamma-tocopherol; forage plant; ds.
 XX Synchocystis sp.
 OS
 FH Key Location/Qualifiers
 FT CDS 1..954
 FT /*tag= a
 FT /gene= "SLR0089"
 FT /product= "gamma-TMT"
 PN MO9904622-A1.
 PD 04-FEB-1999.
 XX
 PF 22-JUL-1998: 98MO-US15137.
 XX
 PR 17-JUL-1998: 98US-0053819.
 PR 25-JUL-1997: 97US-0053819.
 PR 26-JAN-1998: 98US-0072497.
 XX
 PA (UYNE-) UNIV NEVADA.
 PI Dellapenna D, Shintani DK:
 XX
 DR MPI: 1999-142458/12.
 DR P-PSDB: AAM95016.
 XX
 PT Newly isolated DNA fragment comprising a ^c-tocopherol (vitamin E)
 PT methyltransferase coding sequence - useful for producing
 PT tocopherol profile
 XX
 PS Claim 2: Page 33-35; 46pp; English.
 XX
 CC The invention provides DNA sequences encoding gamma-tocopherol
 CC methyltransferases (gamma-TMT). The gamma-TMT encoding genes SLR0089
 CC and 165H5T7 are isolated from Synchocystis and Arabidopsis species
 CC respectively. The DNA fragments are useful for producing transgenic
 CC plants containing a higher alpha-tocopherol:gamma-tocopherol ratio. They
 CC are also useful for producing alpha-tocopherol and plants with increased
 CC gamma-tocopherol, which may be useful in certain industries such as the
 CC meat industry e.g. for developing forage plants to feed animals. The
 CC production of transgenic plants (and seeds) with a higher alpha-
 CC tocopherol:gamma-tocopherol ratio, using the gamma TMTs, will increase
 CC the level of alpha-tocopherol in the human diet, in addition to enhancing
 CC the stability and shelf life of plants and plant products. Increased
 CC levels of alpha-tocopherol will also increase meat quality and extend
 CC shelf life of post-processed meat products. Plants with a higher alpha-
 CC tocopherol:gamma-tocopherol ratio may also produce advantageous
 CC phenotypes. The present sequence represents the SLR0089 gene encoding a
 CC Synchocystis gamma-TMT.
 XX
 SQ Sequence 954 BP; 204 A; 232 C; 271 G; 247 T; 0 other;
 Alignment Scores:
 P-Id. No.: 1.31e-57 Length: 954
 Score: 647.50 Matches: 135
 Percent Similarity: 63.16% Conservative: 45
 Best Local Similarity: 47.37% Mismatches: 102
 Query Match: 35.38% Indels: 3
 DB: 20 Gaps: 3
 US-09-857-613a-28 (1-350) x AAX17788 (1-954)
 OY 65 LeuGLNLYSGLYIleAlaIuHeTYrAspGluSerSergLYLeuTrpLusnIleTrp 84
 DB 100 CTCTACGAAAAAATTAAATTTCTACGACGACTCCAGCGGTCTCGGAAGAAGCTTGG 159
 OY 85 GLAspHISMetHISGLYpHeTYrAspSerAspSerThValSerLeuSerAspHIS 104
 DB 160 GGTGAGCATATGCAACGAGGCTACTACGCTCCCAAGGACCTATTCGATC---GATGCC 216

OY 105 ArgAlaAlaGlnIleArgMetIleGlnIuSerLeuAlaGpHeAlaSerValSerGlu 124
 DB 217 CCGCAGGCTCAATATGATTCATCAAAAGAACATATTGGCTGGCA---GTGCCCAAAAT 273
 OY 125 ArgSerLYSTrProlYSerIleValAspValGlyGlyIleGlyIleSerArg 144
 DB 274 AGCGCCAAA---CCACGAAAAATTTCTCATTTAGGCTGTGCATTGGCGGCAAGATTGG 330
 OY 145 TytleuAlaLYSLysPheGlyAlaThrSerValGlyIleThrLeuSerProValGlnAla 164
 DB 331 TACTTGGCCAGCAACCAACCAAGAGTAGTGGGGCTACTCTTCCCGACGTCAGGTCG 390
 OY 165 GlnArgAlaAsnAlaLeuAlaAlaAlaGlnGlyLeuAlaAspLYSLysPheGlnVal 184
 DB 391 GAACGGGGGGGCAACGGCCAGGGCCCTGGGCTTCACCTCCAGTTTCAGGTG 450
 OY 185 AlaAspAlaLeuGlnGlnProPheSerAspGlyGlnPheAspLeuValITrPserMetGlu 204
 DB 451 GCCAATGCTTGGATTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGAA 510
 OY 205 SerGlyGlnHISMetProAspLYSLysAlaLYSLysPheValGlyLeuAlaArgValAla 224
 DB 511 AGTGGGAGCAGCATGCCCAACAAAGCTCACTTTTACACAGACTTGGCGGCTTAA 570
 OY 225 ProGlyAlaIleIleIleIleValIThrTrpCysHISArgAspLeuGlyProAspGln 244
 DB 571 CCAGGTGGCCGCTGATTTTATGCGACCTGTGTCATCTGCCATTCATGTCCTCCGCAATGGC 630
 OY 245 SerLeuHISProTrpGlnAspLeuLYSLysIleGlyAspAlaLYSLysPro 264
 DB 631 CCCGTACTGCGGATGAAGCGCCATTCACAGCCATTCATGACCTTACTGCTTGGCC 690
 OY 265 AlaTrpCysSerThrSerAspTrpValLYSLysLeuGlnSerLeuSerLeuGlnAspIle 284
 DB 691 TATGTGTTTCCCTCCGACCTACGAGCGATGCCAGGATGTGGCTTGGCAAAAT 750
 OY 285 LysSerGluAspTrpSerAlaPheValAlaProPheTrpProlAlaValIleArgSerAla 304
 DB 751 AAGATGCGGATGTCTAGTGGCGGTGGCCACTTTTGGAGCGGCTATGATTCGCTGCG 810
 OY 305 PheThrTrpLYSLysLeuSerSerLeuSerSerGlyGlnLYSLysIleGlyAla 324
 DB 811 TTGCATCCCGGCTGTGGCTTGGCGCAAGCGGGCCAAAATTAATCAATGCCGCC 870
 OY 325 LeuAlaMetProLeuMetIleGlnIleLYSLysAspLeuIleLYSLysPheAlaIleIle 344
 DB 871 CTGTGTTTACGATTAATGAATGGGCTATGACAGCGGATTAAGTCGTTGGCTTATTA 930
 OY 345 ThrCysArgLYSPro 349
 DB 931 ACGGCGATTAAGCCT 945
 RESULT 14
 AAF26173
 ID AAF26173 standard; DNA; 930 BP.
 XX
 AC AAF26173:
 XX
 DT 26-APR-2001 (first entry)
 XX
 DE Synchocystis PCC6803 MPWT encoding DNA SEQ ID 7.
 XX
 KW 2-methyl-6-phytylhydroquinone methyltransferase; MPWT: transgenic plant;
 KW tocopherol; tocotrienol; vitamin E; 2,3-dimethyl-6-phytyl hydroquinone;
 KW 2,3-dimethyl-6-geranylgeranyl hydroquinone; ds.
 OS Synchocystis sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 10..915
 FT /*tag= a
 FT /product= "MPWT"

Result No.	Score	8			ID	Description
		Query Match	Length	DB		
1	392	21.4	957	4	US-09-382-906A-1	Sequence 1, Appl1
2	261	14.3	2122	4	US-09-029-603-1	Sequence 1, Appl1
3	225	12.3	6085	4	US-09-029-603-4	Sequence 4, Appl1
4	220.5	12.0	80161	3	US-09-036-987A-1	Sequence 1, Appl1
5	220.5	12.0	80161	4	US-09-370-700-1	Sequence 1, Appl1
6	189	10.3	1400	4	US-09-041-718-1	Sequence 1, Appl1
7	185	10.1	49377	1	US-08-764-223A-1	Sequence 1, Appl1
8	156.5	8.6	1149	2	US-08-844-305-1	Sequence 1, Appl1
9	152.5	8.3	80161	3	US-09-036-987A-1	Sequence 1, Appl1
c 10	152.5	8.3	80161	4	US-09-370-700-1	Sequence 1, Appl1
11	138	7.5	4833	4	US-09-066-047-1	Sequence 1, Appl1
12	137.5	7.5	3236	4	US-08-961-527-222	Sequence 222, Appl1

13	131	7.2	8051	2	US-08-576-656A-2	Sequence 2, Appl 1
14	127	6.9	4403765	4	US-09-103-840A-2	Sequence 2, Appl 1
15	120	6.6	3719	1	US-08-920-812-10	Sequence 10, Appl
16	120	6.6	3719	1	US-08-920-827-10	Sequence 10, Appl
17	120	6.6	3719	1	US-08-921-177-10	Sequence 10, Appl
18	120	6.6	3719	1	US-08-362-577C-10	Sequence 10, Appl
19	120	6.6	3719	2	US-08-920-828-10	Sequence 10, Appl
20	117.5	6.4	4411529	4	US-09-103-840A-1	Sequence 1, Appl
21	114.5	6.3	2451	3	US-09-217-609A-21	Sequence 21, Appl
22	114.5	6.3	2451	4	US-08-873-235B-21	Sequence 21, Appl
23	114.5	6.3	4403765	4	US-09-103-840A-2	Sequence 2, Appl
24	113.5	6.2	741	3	US-09-217-609A-29	Sequence 29, Appl
25	113.5	6.2	741	4	US-08-873-235B-29	Sequence 29, Appl
26	109	6.0	9377	4	US-09-221-017B-1002	Sequence 1002, Appl
27	108.5	5.9	2067	4	US-08-760-489-1	Sequence 821, App
28	108.5	5.9	4463	2	US-08-185-373-1	Sequence 1, Appl
29	108.5	5.9	4463	2	US-08-185-373-3	Sequence 3, Appl
30	108.5	5.9	4463	4	US-09-185-373-1	Sequence 1, Appl
31	108.5	5.9	4463	4	US-09-481-068-11	Sequence 65, Appl
32	107.5	5.9	8257	4	US-08-404-970B-65	Sequence 11, Appl
33	104	5.7	756	2	US-08-846-338-11	Sequence 11, Appl
34	104	5.7	756	2	US-08-411-768B-1	Sequence 1, Appl
35	104	5.7	5872	3	US-08-411-768B-6	Sequence 6, Appl
36	104	5.7	5872	3	US-09-718-692-1	Sequence 1, Appl
37	103	5.6	4911	4	US-09-718-852-1	Sequence 1, Appl
38	103	5.6	4911	4	US-09-718-852-1	Sequence 1, Appl
39	103	5.6	4911	4	US-08-457-245-1	Sequence 1, Appl
40	102.5	5.6	5100	1	US-09-103-840A-1	Sequence 4, Appl
41	101.5	5.5	5822	3	US-08-899-595-5	Sequence 5, Appl
42	101.5	5.5	5822	3	US-08-936-165A-119	Sequence 119, App
43	101.5	5.5	1406	4	US-09-134-001C-407	Sequence 407, App
44	99.5	5.4	777	4		

ALIGNMENTS

US-09-382-906A-1

Sequence 1 Application US/09382906A
 Patent No. 6448475
 GENERAL INFORMATION:
 APPLICANT: Dellapenna, Dean
 APPLICANT: Shiltan, David
 TITLE OF INVENTION: Manipulation of Tocopherol Levels in Transgenic Plants
 FILE REFERENCE: 920905.90032
 CURRENT APPLICATION NUMBER: US/09/382,906A
 CURRENT FILING DATE: 1999-08-25
 PRIOR APPLICATION NUMBER: 60/097,863
 PRIOR FILING DATE: 1998-08-25
 NUMBER OF SEQ ID NOS: 4
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 1
 LENGTH: 957
 TYPE: DNA
 ORGANISM: *Synechocystis* PCC6803
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)..(954)
 US-09-382-906A-1

Alignment Scores:

Pred. No.:	6,31e-37	Length:	95
Score:	392.00	Matches:	92
Percent Similarity:	51.76%	Conservative:	40
Best Local Similarity:	36.08%	Mismatches:	37
Query Match:	21.42%	Indels:	86
OB:	4	Gaps:	10

US-09-857-613A-28 (1-350) x US-09-382-906A-1 (1-957)

69 ILeaIaGluphetyraspGlu-----SerserGlyLeuTrpGluAsnIleTrpGlyasp 86


```

Sequence 4 Application US/09/29603
Patient No. 6210935
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Engel, Natalie
APPLICANT: Bietenhader, Jung
APPLICANT: Toupet, Christine
APPLICANT: Pospisich, Andreas
TITLE OF INVENTION: Staurosporin Biosynthesis Gene Clusters
FILE REFERENCE: 4-20555/A/PCT
CURRENT APPLICATION NUMBER: US/09/029,603
CURRENT FILING DATE: 1998-03-20
EARLIER APPLICATION NUMBER: PCT/EP96/03643
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 6085
TYPE: DNA
ORGANISM: Streptomyces longisporoflavus
FEATURE:
NAME/KEY: misc_RNA
LOCATION: (378)..(1665)
OTHER INFORMATION: ORF
FEATURE:
NAME/KEY: misc_RNA
LOCATION: (1747)..(2553)
OTHER INFORMATION: ORF
FEATURE:
NAME/KEY: misc_RNA
LOCATION: (2593)..(4011)
OTHER INFORMATION: ORF
FEATURE:
NAME/KEY: misc_RNA
LOCATION: (4013)..(4999)
OTHER INFORMATION: ORF
FEATURE:
NAME/KEY: misc_RNA
LOCATION: (5071)..(6085)
OTHER INFORMATION: ORF
US-09-029-603-4

Alignment Scores:
Pred. No.: 7,93e-16 Length: 6085
Score: 225.00 Matches: 75
Percent Similarity: 41.45% Conservative: 51
Best local Similarity: 24.67% Mismatches: 120
Query Match: 12.30% Indels: 58
DB: 4 Gaps: 12

US-09-857-613A-28 (1-350) x US-09-029-603-4 (1-6085)
QY 33 ProargserTrpAlaProIleatrgAlaSerAlaalaSerSer-----GluArgGlyGlu 50
||| ||| |||::: ||| :::: |||
Db 1684 CCCTTGGCCCTCGGCCACCCCTTCAGACCAGCCCGCCTCTGTGACTCCGACAGCAGGAAGGCACA 1743
|||||:::|||||
QY 51 lIleValLeuglucInuLySProLySLySAspAspLySLySLySleuGlInuLySlYLleAla 70
::: ||| |||
Db 1744 TCCATGACGACGAGCATGCCACACCCAGCCGCCGAC-----TCGGTCGGT 1785
|||||:::|||||
QY 71 GluPheTyrraspGlu-----SerSerGlyLeuTrpGluasnIlleTrpGlyAspHisMet 88
||| ||| |||::: ||| :::: |||
Db 1786 GAGGTGACGACGACGATTGCCGCGCACGCCGCGGCC---AGCACCGCATGGCGCGCAACATC 1842
|||||:::|||||
QY 89 HisrIsGlyPheTyrrAspSerAspSerThrValSerLeuSerAspHisarGlnlaIn 106
||| |||:::||||| ||| :::: |||
Db 1843 CACGTGGGGTGACTGGGACGACGACCCCCGAGGTGCGATGCGCGAAGCCACGACCGGCTC 1902
|||||:::||||| |||
QY 109 lIleaRgmetlIeGluInuSerLeuArGpheaAlaSerValSerCluGluArgSerTyStrp 128
::::: ||| |||
Db 1903 ACCGATCTCGTGGCGGACGCCCTCGGGGCTC-----CGCCCC 1938
::::: ||| |||
QY 129 ProLysserlIleValaspaValGlyCySgglYllieGlylYSerSerArGtyrlleuAlaLyS 148

```

```

Db 1939 GACCGGATCTGCTGGACGCGGGCTCGCGCATCGGCGTGGCGGCTCTTCAGATGACCGCGA 1998
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 149 LysPheGlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAsn 168
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1999 GCGGACGACGCTCGCGCTCACCGGGATCACCGTACACGACAGCAGTCCAGAGCGGCC 2058
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 169 AlaLeuAlaIalaGlnGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAlaLeu 188
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2059 GAGCGGGGGTGGAGTCCGATCCGATCCGGGGCGGGCGGGCTCTCCCTCGGCTCGCGGACCCCATG 2118
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 189 GlnGlnProPheSerAspGlyGlnPheAspLeuValITripSerMetGlnSerGlyGlnIle 208
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2119 GACCTCCCTTCGAGAGACGCTCTTCGACAGCGCGCTTGCCATCGATGAGTCCGTGTCAT 2178
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 209 MetProAspLysAlaLysPheValGlyGlnLeuAlaArgValAlaAlaProGlyValIle 228
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2179 CTGCGCGACGACACACCGCGGCTCAGAGAGATCCACCGGATCGTCCGCCCGCGCGCGCGCG 2238
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 229 IleIleIleValIleThrTrpCysHisArgAsp-----LeuGlyProAspGlnGlnSerLeu 246
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2239 CTCGTCAATCCGACCTGTGTACGACAGACGCCGTTCACCGGGCGCGACAGAGAGGTGCTC 2298
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 247 His-----ProITripGlnGlnAspLeu 253
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2299 GACGGATGCTGCTGATGTACAGAGATCGCGCGGATCCACACACCTTACGAGAT----- 2352
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 254 LeuLysLysIleCysAspAlaArgTyrIleuProAlaITripCysSerThrSerAspTyrVal 273
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2353 CGCGCGGACGTGGCGGCGGCG-----GGCTG----- 2379
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 274 LysLeuLeuGlnSerLeuSerLeuGlnAspIleLysSerGlnAspITripSerArgPheVal 293
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2380 -----GAACTGCTCGGAGCGACGAGCATCGGTGACAG-----GTC 2415
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 294 AlaProHeITripProAlaValIleArgSerAlaPheITripLysGlyLeuSerSerLeu 313
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2416 CGCGCTACTACGCG-----CATGCCGCGCGCGCGCTTCGGGGGTCTCCGCGGGCT 2466
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 314 LeuSerSerGly 317
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2467 CTCGACCGCGCG 2478
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
US-09-036-987A-1
: Sequence 1, Application US/09036987A
: Patent No. 6143526
:
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Merlo, Donald J.
APPLICANT: Treadway, Patti J.
APPLICANT: Turner, Jan R.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
TITLE OF INVENTION: Production
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences LLC Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,987A

```


Db 1198 GGTGGTACGCAAAATTTCTTACTCCATGCATGATTGTCGTTTAGAACA 1231

332 GlucyltyrtylsyaspIlelleysphalallethnrcysArgLysPro 349
111 :::111 :
111111111

RESULT 7
05-08-764-233A-1
Sequence 1, Application US/08764233A
Patent No. 5716849

GENERAL INFORMATION:
APPLICANT: Ligon, James M.
APPLICANT: Schupp, Thomas
APPLICANT: Beck, James J.
APPLICANT: Hill, Dwight S.
APPLICANT: Neff, Snezana
APPLICANT: Ryals, John A.

TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,233A
FILING DATE:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/729,214
FILING DATE: 09-OCT-1996
PRIOR APPLICATION DATA: US 08/258,261
FILING DATE: 08-JUN-1994

ATTORNEY/AGENT INFORMATION:
NAME: Melgs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: 1506/CIP6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 49377 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:
ORGANISM: Sorangium cellulosum
IMMEDIATE SOURCE:
CLONE: p98/1, pUJ3, and pVKM15

FEATURE:
NAME/KEY: misc.feature
LOCATION: 383..760

OTHER INFORMATION: /product= "Sora"
/note= "this gene encodes a protein that is highly homologous
to type I PKs such as eryA from
Saccharopolyspora erythraea."

OTHER INFORMATION: are known to be involved in the synthesis of polyketide
compounds."

FEATURE:
NAME/KEY: misc.feature
LOCATION: 927..19874

OTHER INFORMATION: /product= "Sora"
/note= "Gene product is highly homologous to type I PKs that
are known to be involved in the synthesis of polyketide
compounds."

FEATURE:
NAME/KEY: misc.feature

```

LOCATION: 942..7115
OTHER INFORMATION: /product= "Module 1 of SorA"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7203..12884
OTHER INFORMATION: /product= "Module 2 of SorA"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 13455..19616
OTHER INFORMATION: /product= "Module 3 of SorA"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19871..46318
OTHER INFORMATION: /product= "SorB"
OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKs ge
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19870..24556
OTHER INFORMATION: /product= "Module 1 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 24638..30820
OTHER INFORMATION: /product= "Module 2 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 30881..35446
OTHER INFORMATION: /product= "Module 3 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 35528..40114
OTHER INFORMATION: /product= "Module 4 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 40190..46318
OTHER INFORMATION: /product= "Module 5 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 46851..47891
OTHER INFORMATION: /product= "Sorm"
OTHER INFORMATION: /note= "The protein encoded by the sorm gene is highly
OTHER INFORMATION: homologous to the methyltransferase from Streptomyces
OTHER INFORMATION: hygroscopicus that is involved in the synthesis of the
OTHER INFORMATION: polyketide rapamycin."
US-08-764-233A-1
Alignment Scores:
Pred. No.: 1.28e-09 Length: 49377
Score: 185.00 Matches: 89
Percent Similarity: 39.788 Conservative: 57
Best Local Similarity: 24.258 Mismatches: 129
Query Match: 10.11% Indels: 92
DB: 1 Gaps: 16
US-09-857-613A-28 (1-350) x US-08-764-233A-1 (1-49377)
QY 17 ThrPheArgSerGlnSerProArgThrPheAlaArgIleatGValGlyProArgSerTrp 36
:::|||||::: ||| 111
46821 AGCTTCGCGTGGGGCGCATGAAGACGCCCATGCAAAATTTCGCCAGGGGCTCGGACA--- 46877
-----ArgAlaSerAlaAlaSerSerGlnArgGlyGluLe 51
37 AlaProIle:::||||| 111 111
46878 TCTCCCATCCAGCGCGGATAGCAACAGAGAGCAATATGCGATCCACACAGATCGACGTG 46937
-----ValleuGluGlnIlyProLysLysAspAspLysLysLysLeuGlnIlySly 68
52 ValleuGluGlnIlyProLysLysAspAspLysLysLysLeuGlnIlySly 68
DB 46938 ATCAAAATGTTCAACGATGCTGGAGCGCCATGGCCAGCATCCAGTCGCTCGACCG 46997
::: 111
69 -----IleAlaGluPheTyrAsp 74
QY 46998 GACGCGCTCGCGCTCAAGGTGGACAGCTCAACGCCCTCATGAGAGCGGCTAACAACCTC 47057
::: 111
75 -----GluSerSerGlyLeuTrpGluAsnIleTrpGlyAspHisMetHisSgly 91
::: 111 111 111

```

```

Db 47058 GAGCCGCTCAAGTCCGACGCTTTCACAAAC---TGGGGG-----ATGTTTCAGCAG 47108
QY 92 PheTYrAspSerSerThrValSerLeuSerAspHisArgAlaIleIleArgMet 111
Db 47109 GAGGTTTTCGTGATGTTCTTCAACAGCTGCCAGATTACGACAAAGCAGCAGCAGCAGT 47168
QY 112 IleGlnIle-----SerLeuArgPheAlaSerValSerGlnIleArgSer 126
Db 47169 TATTGAGACAGCTCTACGCTACAGCGTGAAGAGTTCCGGCCGAGAGAGATCG--- 47225
QY 127 LysTrpProLysSerIleValAlaAspValGlyCysGlyIleGlyIleSerArgTrpIleu 146
Db 47226 -----CCAGCCAGATCTCGAGCTCGCAACACTGCGAAGGCGCTCAACTCTCT 47279
QY 147 AlaLysLysPheGly-----AlaThrSerValGlyIleThrLeuSerProValGlnIle 165
Db 47280 TCCCGGATCGAGGCTCGCAGACACCTTCTCGGCTCGACCTGTCCAGCAGCAGGCTGAC 47339
QY 166 ArgAlaAsnAlaLeuAlaAlaIleGlnIleLeuAlaAspLysValSerPheGlnValAla 185
Db 47340 ATCGCCAAAGCCCGGCTTCTCGCAGACCGGGC-----TCGTTGACCTACGTTCAAGGC 47390
QY 186 AspAlaLeuGlnIleProPheSerSerAspGlyIlePheAspLeuValTrpSerMetGlnSer 205
Db 47391 GACGGGAGAACTCTTCTTCCGCGCAGCGCATTCGACGCTGTCATCATGTCAGAGAGC 47450
QY 206 GlyIleHisMetProAspLysAlaLysPheValGlyIleLeuAlaArgValAlaIlePro 225
Db 47451 TCGCAACAACCTCAATCTCGCGCAATCATCTCGAAGTGGCGGCTGCTGCGACCG 47510
QY 226 GlyAlaIleIleIleIleValThrTrpCysHisArgAspLeuGlyIleProAspGlnIleSer 245
Db 47511 GCGCGGCTTC-----TTTTCGATGTGATGTCTCTCGCAACAATGCGC--- 47552
QY 246 LeuHisProTrpGlnIleAspLeuLeuLysIleCysAspAlaTrpTrpIleuProAla 265
Db 47553 -----TATTCGGTGTATGCGAG 47567
QY 266 TrpCys-----SerThrSerAspTrpValLysLeuGlnIleSerLeuSerLeuGlnAsp 283
Db 47568 AATTCCAAAGCAGCAGACCCAGCGGCGCAGCTCGACTGCTCAAGAGAACGACACATCTCGGAG 47627
QY 284 IleLysSerGlnAspTrpSerArgPheValAlaProPheTrpProAlaValIleArgSer 303
Db 47628 TACGTAAGAAGAGCTATCCGCGCAGCAGCTGCGCGCGC----- 47666
QY 304 AlaPheThrTrpLysGlyLeuSerSerLeuLeuSerSerGlyGlnLysThrIleLysGly 323
Db 47667 -----ACCAAGCGCAAGCAGCAGCTGCGAGCGC 47693
QY 324 AlaLeuAlaMetProLeu-----MetIleGlnIleGlyTrpLysLys 336
Db 47694 GCGCTGCGCTACCCGCTCGGAAGCTTTTCAGCTCGGTGATGATGATGCGGCTACGAGATCC 47753
QY 337 AspLeuIleLysPheAlaIle 343
Db 47754 GAG-----TTTCCGCTC 47765

RESULT 8
; Sequence 1, Application US/08844305
; Patent No. 5936139
; GENERAL INFORMATION:
; APPLICANT: Schmid, Katherine M.
; TITLE OF INVENTION: Cyclopropane Fatty Acid Expression in Plants
; FILE REFERENCE: M5-403F
; CURRENT APPLICATION NUMBER: US/08/844,305
; CURRENT FILING DATE: 1997-04-10
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1149
; TYPE: DNA

```

```

; ORGANISM: Escherichia coli
; US-08-844-305-1

Alignment Scores:
Pred. No.: 7,18e-09 Length: 1149
Score: 156.50 Matches: 69
Percent Similarity: 41.45% Conservative: 45
Best Local Similarity: 25.09% Mismatches: 105
Query Match: 8.55% Indels: 56
DB: 2 Gaps: 14

US-09-857-613a-28 (1-350) x US-08-844-305-1 (1-1149)

QY 71 GluPheTYrAspGlnSerSerGlyLeuTrpGlnAsnIleTrpGlyAspHisMetHis 90
Db 352 GAGCATTTACGATTTGGGTATGATGATCTTGTCCAGCCGATCTTGATCCCTTCAATGAT 411
QY 91 Gly-----PheTYrAspSerSerThrValSerLeuSerAspHisArgAlaAla 107
Db 412 TCTGCGCTTACTGCAAGATGCCGAT-----AATCTGGAATCTGCCAGCAGCGC 462
QY 108 GlnIleArgMetIleGlnIleSerLeuArgPheAlaSerValSerGlnIleArgSerLys 127
Db 463 AAGCTCAAAATGATTTGTAAAAATTCAGCTTA-----AAA 498
QY 128 TrpProLysSerIleValAspValGlyCysGlyIleGlyIleSerSerArgTrpIleuAla 147
Db 499 CCAGGATCGCGCTACTGATATATGCTGCTGCGGCTGGGCGGAGCTGCACTACATGATGCA 558
QY 148 LysLysPheGlyAlaThrSerValGlyIleThrLeuSerProValGlnIleArgAla 167
Db 559 TCTAATATTAGACGTAAAGCTGCGGCGCTCACCATTTCTGCCAAGCAGCAAAAATGCGCT 618
QY 168 AsnAlaLeuAlaAlaIleGlnIleLeuAlaAspLysValSerPheGlnValAlaAspAla 187
Db 619 CAG-----GAACGCTGTAAAGCCTGATGTCCACATT--- 651
QY 188 LeuGlnIleProPheSerSer-----GlyIlePheAspLeuValTrpSerMetGlnSer 205
Db 652 TTGCTCCAAAGATTATGCTGATGACCTGCAAGCAGCACTTGATGATGCTTCTTCTGGGAGAG 711
QY 206 GlyIleHisMetProAspLys-----AlaLysPheValGlyIleLeuAlaArgValAla 223
Db 712 TTCGACGACGTGCGAGCAAAATTCAGATACCTATTGCGCGTGGATGATGATTTG 771
QY 224 AlaProGlyAlaIleIleIleIleValThrTrpCysHisArgAspLeuGlnProAspGly 243
Db 772 AAACCGAAGGCAATATTCCTGCTCCATCT-----ATCGGTTCCGAAAAA 816
QY 244 GlnSerLeuHis-----ProTrpGlnIleAspLeuLeuLysIleCysAspAlaTrp 261
Db 817 ACCGATCTGAATGTTGATCCCTCG-----ATTATTAATAT 852
QY 262 TYrLeuProAlaTrpCysSerThrSerAspTrpValLysLeuGlnIleSerLeuSerLeu 281
Db 853 ATTTTCCGACAGCGTTCGCCCTCT-----GTACGCCAGATTCTCAGTCAGCGAA 906
QY 282 GlnAspIleLysSerGlnAspTrpSerArgPheValAlaProPheTrpProAlaValIle 301
Db 907 CCCCACTTTGTGATGAGACACTGCGATTAACCTTCGCTGATTAAC----- 951
QY 302 ArgSerAlaPheThr--TrpLysGlyLeuSerSerLeuLeuSerSerGlyGlnLys----- 319
Db 952 ---GATACTACGTTGATGCGGTGTATGAGACATTCCTCGCCGATGCGCCAGAAATTCGG 1008
QY 320 ---ThrIleLysGlyAlaLeu-----AlaMetProLeuMet 330
Db 1009 GATAACTATAGTAACGCTTTAAACGAATGTTTACTATTATTC 1051

RESULT 9
; Sequence 1, Application US/09036987A
; Patent No. 6143526

```


GENERAL INFORMATION:

APPLICANT: Baltz, Richard H.

APPLICANT: Broughton, Mary C.

APPLICANT: Crawford, Kathryn P.

APPLICANT: Madduri, Krishnamurthy

APPLICANT: Merlo, Donald J.

APPLICANT: Treadway, Patti J.

APPLICANT: Turner, Jan R.

APPLICANT: Maldon, Clive

TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide

TITLE OF INVENTION: Production

NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dow Agrosciences LLC Patent Department

STREET: 9330 Zionsville Road

CITY: Indianapolis

STATE: Indiana

COUNTRY: USA

ZIP: 46268

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER: US/09/036,987A

FILING DATE: 09-MAR-1998

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Stuart, Donald R.

REGISTRATION NUMBER: 28,479

REFERENCE/DOCKET NUMBER: 50,608

TELECOMMUNICATION INFORMATION:

TELEPHONE: (317)337-4816

TELEFAX: (317)337-4847

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 80161 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-09-036-987A-1

Alignment Scores:

Pred. No.: 2e-05 Length: 80161

Score: 152.50 Matches: 48

Percent Similarity: 38.38% Conservative: 28

Best Local Similarity: 24.24% Mismatches: 77

Query Match: 8.33% Indels: 45

DB: 3 Gaps: 4

US-09-857-613A-28 (1-350) x US-09-036-987A-1 (1-80161)

QY 85 GtYAspHsmEhIsGlypHeTyRAsp----- 94

DB 13517 GGGCGTCCCTCCACCGAGTACTGGGCGGGGTATCGGAGATCCCGTCCACA 13458

QY 95 -----SerAspSerThValSerLeuSerAspHisArgAlaAlaGlnIleArgMetIle 112

DB 13457 CCGTGTCCGATGCTGCCAGCACTGACCGAC-----CTGTTCAATCGAC 13413

QY 113 GlnGluSerLeuArgPheAlaSerValSerGluGluArgSerLysTrpProLysSerIle 132

DB 13412 AAGCGCGCGCTCCGTCGCGAGCG-----CACCTG 13383

QY 133 ValAspValGlyGlySerGlyIleGlyGlySerArgTyrLeuAlaLysLysPheGlyAla 152

DB 13382 TTCGACTGGCGTGGCGCAATGGGCGAGCCGTAAGTCCGTGGCGATGCCCGCGCGTT 13323

QY 153 ThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAsnAlaLeuAla 172

DB 13322 CGAGTCAACCGGATCACTGCTGAACGCCGACGATCTCGCGCGCCACGAGGTCGCGAAC 13263

QY 173 AlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAlaLeuGlnGlnPhe 192

DB 13262 GAGACCGGATCGCGCGCATCTTCACTGATCTGACAGCGCGCCACGACTCCCTAC 13203

QY 193 SerAspGlyGlnPheAspLeuValTrpSerMetCysLeuSerGlyLysIleMetProAspLys 212

DB 13202 CCGGACGCTTCTTTCAGCGCGCATGCGCATGCTCGTGCAGATCTGTCGACAG 13143

QY 213 AlaLysPheValGlyGluLeuAlaArgValAlaAlaLapProGlyAlaIleIleLeuVal 232

DB 13142 GCCGCGCGATCGCGCGCATCTGACGATCTCGAACCAGCGCGCGTTCCTCGGA 13083

QY 233 ThrTrpCysHisArgAspLeuGlyProAspGluGlnSerLeuHisProTrpGluGlnAsp 252

DB 13082 GACATCATCACTCGGCTTCGATCGATCCCGAAGAG----- 13050

QY 253 LeuLeuLysLysIleCysAspAlaTyrTyrLeuProAlaTrpCysSerThrSer 270

DB 13049 -----TACGCGCGGCTTTCGACGCGCAGCAGC 13023

RESULT 10

US-09-370-700-1/c

Sequence 1, Application US/09370700

Patent No. 6274350

GENERAL INFORMATION:

APPLICANT: Baltz, Richard H.

APPLICANT: Broughton, Mary C.

APPLICANT: Crawford, Kathryn P.

APPLICANT: Madduri, Krishnamurthy

APPLICANT: Turner, Jan R.

APPLICANT: Maldon, Clive

TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide

FILE REFERENCE: 50489 DIV1

CURRENT APPLICATION NUMBER: US/09/370,700

EARLIER FILING DATE: 1999-08-09

EARLIER APPLICATION NUMBER: US 09/36987

EARLIER FILING DATE: 1998-03-09

NUMBER OF SEQ ID NOS: 39

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1

LENGTH: 80161

TYPE: DNA

ORGANISM: Saccharopolyspora spinosa

US-09-370-700-1

Alignment Scores:

Pred. No.: 2e-05 Length: 80161

Score: 152.50 Matches: 48

Percent Similarity: 38.38% Conservative: 28

Best Local Similarity: 24.24% Mismatches: 77

Query Match: 8.33% Indels: 45

DB: 4 Gaps: 4

US-09-857-613A-28 (1-350) x US-09-370-700-1 (1-80161)

QY 85 GtYAspHsmEhIsGlypHeTyRAsp----- 94

DB 13517 GGGCGTCCCTCCACCGAGTACTGGGCGGGGTATCGGAGATCCCGTCCACA 13458

QY 95 -----SerAspSerThValSerLeuSerAspHisArgAlaAlaGlnIleArgMetIle 112

DB 13457 CCGTGTCCGATGCTGCCAGCACTGACCGAC-----CTGTTCAATCGAC 13413

QY 113 GlnGluSerLeuArgPheAlaSerValSerGluGluArgSerLysTrpProLysSerIle 132

DB 13412 AAGCGCGCGCTCCGTCGCGAGCG-----CACCTG 13383

QY 133 ValAspValGlyGlySerGlyIleGlyGlySerArgTyrLeuAlaLysLysPheGlyAla 152

DB 13382 TTCGACTGGCGTGGCGCAATGGGCGAGCCGTAAGTCCGTGGCGATGCCCGCGCGTT 13323

QY 153 ThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAsnAlaLeuAla 172
 Db 13322 CGAGTACCGGAAATCACCGTAAAGCCAGCATCTCCGCCGCCACAGGCTGCCAAC 13263
 QY 173 AlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAlaAspAlaLeuGlnProPhe 192
 Db 13262 GAGACCGGACATCGCGGAGCTTGTAGTTCATGTCGACAGCGGCGCCAGCTGCCCTAC 13203
 QY 193 SerAspGlyGlnPheAspLeuValTrpSerMetGluSerGlyGlnHisMetProAspLys 212
 Db 13202 CCGGACGCTTCTTTCACGCGCCAGTGGCGCATGCTCGTCCAGATCGTGGACAG 13143
 QY 213 AlaLysPheValGlyGlnLeuAlaArgValAlaAlaProGlyAlaIleIleIleVal 232
 Db 13142 GCCCGCCGATCCCGGAGCTCCACCGAATCTCGAACCAGCGCGGCTGCTCCGGA 13083
 QY 233 ThrTrpCysHisArgAspLeuGlyProAspGlnGlnSerLeuHisProTrpGluInasp 252
 Db 13082 GACATCATCATCGCGGCTCGCATCGCGGAGAG----- 13050
 QY 253 LeuLeuLysLysIleCysAspAlaTyrTrpLeuProAlaTrpCysSerThrser 270
 Db 13049 -----TACCGCGCGGTTTGACGCGCAGCAGC 13023
 RESULT 11
 US-09-066-047-1
 Sequence 1, Application US/09066047A
 Patent No. 6506394
 GENERAL INFORMATION:
 APPLICANT: MURPHY, Cheryl
 BELTZ, Gerald A.
 STOREY, James
 COUGHLIN, Richard T.
 TITLE OF INVENTION: NUCLEIC ACIDS, PROTEINS, AND METHODS OF
 USE OF GRANDLOCYTIC ERHL1CH1A
 NUMBER OF SEQUENCES: 41
 CORRESPONDENCE ADDRESS:
 ADDRESSSEE: HALE AND DORR LLP
 STREET: 60 State Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: United States
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/066,047A
 FILING DATE: 24-Apr-1998
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/044,869
 FILING DATE: 25-Apr-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Superko, Colleen
 REGISTRATION NUMBER: 39,850
 REFERENCE/DOCKET NUMBER: 106,941.156
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 526-6000
 TELEFAX: (617) 526-5000
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 483 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-066-047-1

Alignment Scores:
 Pred. No.: 1,14e-05 Length: 483
 Score: 138.00 Matches: 55
 Percent Similarity: 44.93% Conservative: 38
 Best Local Similarity: 26.57% Mismatches: 83
 Query Match: 7.54% Indels: 31
 DB: 4 Gaps: 8
 US-09-857-613a-28 (1-350) x US-09-066-047-1 (1-4833)
 QY 92 PheTyrAspSerAspSerThrValSerLeuSerAspHisArgAlaAlaGlnIleArgMet 111
 Db 3713 TCGTGAGATGGGAGCTCTTTTCAGCTTTCACAGATGAATATCTTCCGCTTACAGTAT 3772
 QY 112 IleGlnGluSerLeuArgPheAlaSerValSerGluGluArgSerLysTrpProLysSer 131
 Db 3773 ATTCTTGAATAATTTACAAAGGCTACTACTACAGCTAAAG----- 3814
 QY 132 IleValAspValGlyCysGlyIleGlyGlySerSerArgTrpLeuAlaLys--LysPhe 150
 Db 3815 CTTTGGATATGCGTGTGTGTGTGTGTATTTGGCAAGCCATGGCAAGCTTGTGTTT 3874
 QY 151 GlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAsnAlaLeu 170
 Db 3875 ACTGTCACCTGAGATGATCATCATAGAGAAAGAAATGAGACTGCTAGACAGCAGC----- 3928
 QY 171 AlaAlaAlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAlaLeuGln 190
 Db 3929 GCTGCTATGCAAGGCTTATATATAGACCATATTTACGATATAGAGATTTTATATAC 3988
 QY 191 ProPheSerAspGlyGlnPheAspLeuValTrpSerMetGluSerGlyGlnHisMetPro 210
 Db 3989 TCCTCAGAGTGTCTCTTACGATATCATCATCTTATGSAATTTGATGACATATCCCT 4048
 QY 211 AspLysAlaAlaGlnGlyLeuAlaArgValAlaAlaProGlyAlaIleIle 230
 Db 4049 GATTTCAGTAATTTTACTACTCTCTGTAAGTACTGTAACCTGAGGATGCTTTTC 4108
 QY 231 IleValThrTrpCysHisArgAspLeuGlyProAspGlnGlnSerLeuHisProTrpGlu 250
 Db 4109 ATTTCTACA-----CTTAACAGACTATCAATTCATG----- 4141
 QY 251 GlnAspLeuLeuLysLysIleCysAspAlaTyrTrpLeuProAlaTrpCysSerThrse 270
 Db 4142 -----TTACTTGTAAGATAGCTCGCGAATATATCTTC---GCATGGTGC----- 4184
 QY 270 rAspTyrValLysLeuLeuGlnSerLeuSerLeuGlnAspIleLysSerGluAspTrpse 290
 Db 4185 -----CTCCTGGCAGCAGCAGCTGGAAGAAGTTGTCAAGC-----CTTCA 4225
 QY 290 rArgPheValAlaProPhe 296
 Db 4226 GAGATTTCAGATGCCCTAT 4244
 RESULT 12
 US-08-961-527-222
 Sequence 222, Application US/08961527
 Patent No. 6420135
 GENERAL INFORMATION:
 APPLICANT: Charles Kunsch
 TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
 NUMBER OF SEQUENCES: 391
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
 COMPUTER: HP Vectra 486/33


```

QY 26 PheAlaArgIleArgVal-----GlyProArgSerTrpAlaPro 38
    |||||
Db 2213 TTGCGGAGGAGCGCTCTCCGACCCGACGGCAGAGATCGTCCGCGACTGGAGGAC 2272
QY 39 IleArgAlaSerIleAlaSerSerGluArgGlyGluIleValLeuGluGlnLysProLys 58
    |||||
Db 2273 TCACGGCGCCGACGGCGGTACCGGTTTCCGACCGCAAGTCCGTCGACACACACT 2332
QY 59 LysAspArgLysLysLysLeuGlnLysGlyIleAlaGluPheTrpAspGluSerSerGly 78
    |||||
Db 2333 CCGGAGGAGCAGGATGTACGAGGGGGGTTGCGCGAGCTTACGACCGG----- 2383
QY 79 LeuTrpGluAsnIleTrpGlyAspHisMetHisIscGlyPheTrpAspSerSerThr 98
    |||||
Db 2384 -----TTTACCGCGCGCGCGGCGC--- 2401
QY 99 ValSerLeuSerAspHis-----ArgAlaAlaGlnIle---ArgMetIleGlnGluSer 115
    |||||
Db 2402 -----AAGACTACAGCGCGCGCGGAGCGCGGAGCTCCGCGGTGTCAAGACCGC 2452
QY 116 LeuArgPheAlaSerValSerGluGluArgSerLysTrpProLysSerIleValAspVal 135
    |||||
Db 2453 CTCCCTCGGCTTCC-----TCGCTGCTCGACGTG 2482
QY 136 GlyCysGlyIleGlySerSerArgTyrLeuAlaLysLysPheGlyAlaTrpSerVal 155
    |||||
Db 2483 GCCTCGCGGACCGCACCCACCTGCGCGCGCTTCCGACGACGTGAC--- 2539
QY 156 GlyIleThrLeuSerProValGlnAlaGlnAla-----AsnAlaLeuAla 171
    |||||
Db 2540 GGGCTGAGCTGTGCGCGCGGATGATGATGAGTGGCGCGCGCGGACGTGCGCGCATCCG 2599
QY 172 AlaAlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAlaLeuGlnPro 191
    |||||
Db 2600 GTGCTCGAGGCG-----GACATGCGCGACTTGGCGGTG----- 2632
QY 192 PheSerAspGlyGlnPheAspLeuValTrpSerMetGlnSerGly---GluHisMetPro 210
    |||||
Db 2633 -----GATGCGGAGTTCGACCGCGCTGACCTGATTCAGTTCATCGGCGACATGCGC 2686
QY 211 AspLysAlaLysPhe-----ValGlyGluLeuAlaArgValAlaAlaTrpGlyAla 227
    |||||
Db 2687 GAGCGGCGGAGCTGACGACGACGCGCGCTGCTTCCGCGCGACCTCGCGCGCGCGGCGC 2746
QY 228 IleIleIleIleValIleThrTrpCysHisArgAspLeuGlyProAspGluGlnSerLeuHis 247
    |||||
Db 2747 GTGCTGTGTGTCGACCGCTG----- 2767
QY 248 ProTrp---GluGlnAspLeuLeuLysLysIleCysAspAlaTyr---TyrIleuProAlaTr 266
    |||||
Db 2768 ---TGGTTCGCGGAGGACTCTCTC-----GACGCGTACGTCGCGCGGTGACGTG 2812
QY 266 PcySerThrSerAspTyrValLysLeuLeuGlnSerLeuSerLeuGlnAspIleLysSe 286
    |||||
Db 2813 GTCCGCGGAGCAGCTGACGATCTGCGCGTCTGCACTCCGCGCGCGCGCGCGCGCGC 2872
QY 286 rGluAspTrpSerArgPheValAlaProPheTrpProAla-----ValIleArgSerAl 304
    |||||
Db 2873 ACCCGATGTAGATTC-----ACTGGGTGTGTGCGCGCGGTGAACGGTTCGCGCGCAC 2926
QY 304 aPheThrTrpLysGlyLeuSerSerLeuSerSerGlyGlnLysThr 320
    |||||
Db 2927 C---ACGTGAGCAGTACGATCAGCTCTTGTGAGCGGAGAGACG 2972

```

RESULT 14
 ; Sequence 2, Application US/09103840A
 ; Patent No. 6294328
 ; GENERAL INFORMATION:
 ; APPLICANT: FLEISCHMAN, Robert D.
 ; APPLICANT: WHITE, Owen R.
 ; APPLICANT: FRASER, Claire M.

```

; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103, 840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 13.5 Length: 4403765
Score: 127.00 Matches: 55
Percent Similarity: 40.37% Conservative: 33
Best Local Similarity: 25.23% Mismatches: 98
Query Match: 6.94% Indels: 32
DB: 4 Gaps: 5

US-09-857-613a-28 (1-350) x US-09-103-840A-2 (1-4403765)

QY 95 SerAspSerThrValSerLeuSerAspHisArgAlaIleGlnIleArgMetIleGlnGlu 114
    |||||
Db 4172057 GCCGATGAGCAGTCGCGCACACAGATCAAGACGTGTCGCGGACCTATTCGCCGGA 4172116
QY 115 SerLeuArgPheAlaSer----- 120
    |||||
Db 4172117 CATGCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4172176
QY 121 -----ValSerGluGluArgSerLysTrpProLysSerIleValAspValGly 136
    |||||
Db 4172177 TTGGCTGACCACTCGCGCTGAGATGACGAGGAGCGCGCGCGCGCGCGCGCGCGC 4172236
QY 137 CysGlyIleGlySerSerArgTyrLeuAlaLysLysPheGlyAlaTrpSerValGly 156
    |||||
Db 4172237 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4172296
QY 157 IleThrLeuSerProValGlnAlaGlnArgAlaAsnAlaLeuAlaAlaGlnGlyLeu 176
    |||||
Db 4172297 GTGACATCAGGAGATCAACGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 4172356
QY 177 AlaAspLysValSerPheGlnValAlaAspAlaLeuGlnIleProPheSerAspGlyGln 196
    |||||
Db 4172357 ACCGAGCGGCTGCGCTTCCACCTGGGCGGCGCGGATCAAGTCCCGTTCGCGCGACACA 4172416
QY 197 PheAspLeuValTrpSerMetGluSerGlyGluHisMetProAspLysAlaLysPheVal 216
    |||||
Db 4172417 TTGAGCGCGCTGCTGTCAGTGTGCGGCTTCTGCACATTCGCGGACAGAACCGCGCGCC 4172476
QY 217 GlyLeuGlnAlaArgValAlaAlaAlaProGlyAlaIle-----IleIleIleValThrTrp 234
    |||||
Db 4172477 CAGCACTTGTGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4172536
QY 235 CysHisArgAspLeuGlyProAspGluGlnSerLeuHisIleProTrpGluGlnAspLeu 254
    |||||
Db 4172537 GGGGAGCGGCGCGCTGCGCGGCGGAGCGAGCCCATTTGCGCGCGT-----GTC 4172584
QY 255 LysLysIleCysAspAlaTyrTrpLeuProAlaTrpCysSerThrSerAspTyrValLys 274
    |||||
Db 4172585 GCCTGATCGCGCGCGCGCA-----ACCGTACCGCATACACCGAC 4172626
QY 275 LeuLeuGlnSerLeuSerLeu-----GlnAspIleLysSerGluAspTrpSer 290
    |||||
Db 4172627 ATCTCTGAAGGGCGGATTTGCGCACCGCCCATGTGACTCTCATGACGAGAC 4172680

```

RESULT 15

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 19, 2003, 08:06:01 ; Search time 88 Seconds
(without alignments)
3084.279 Million cell updates/sec

Title: US-09-857-613A-28

Perfect score: 1830
Sequence: 1 MATVVRIPITSCIHITFRS.....IEGKKDLKFAITCKRPE 350

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 538826 seqs, 387737923 residues

Total number of hits satisfying chosen parameters: 1077652

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framed_p2n_model -DEV=xlh
-O=/cgn2_1/USPRO.spool/US09857613/funat_12032003.100038.6130/app_query.fasta_1.519
-DB=Published.Applications.NA -OFMT=fastap -SUFFIX=fnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Blosum62
-TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pcr -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFRMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09857613 -ECGN_1_1.53 -funat_12032003.100038.6130
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published.Applications.NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PC1_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PC105_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	386.5	21.1	252	10	US-09-878-574-695
2	316	17.3	311	10	US-09-294-093B-3472
3	226	12.3	1080	9	US-09-938-842A-1815
4	226	12.3	1497	10	US-09-779-144A-6

5	212.5	11.6	1086	9	US-09-938-842A-1856	Sequence 1856, Ap
6	203.5	11.1	1152	10	US-09-801-368-95	Sequence 95, Appl
7	203.5	11.1	1320	10	US-09-779-144A-1	Sequence 1, Appl
8	203.5	11.1	1320	10	US-09-779-144A-5	Sequence 5, Appl
9	196.5	10.7	1420	10	US-09-779-144A-3	Sequence 3, Appl
10	156.5	8.6	3900	10	US-09-805-681-1	Sequence 1, Appl
11	148.5	8.1	411	10	US-09-878-574-3629	Sequence 3629, Ap
12	141	7.7	753	9	US-10-260-877-53	Sequence 53, Appl
13	140	7.7	1801	10	US-09-845-248-1	Sequence 1, Appl
14	139	7.6	1353	10	US-09-738-626-636	Sequence 636, Ap
15	128	7.0	290	10	US-09-294-093B-5022	Sequence 5022, Ap
16	123	6.7	891	9	US-09-938-842A-1172	Sequence 1172, Ap
17	117.5	6.4	618	9	US-09-738-626-1423	Sequence 1423, Ap
18	114	6.2	261	10	US-09-923-876-984	Sequence 984, Ap
19	113	6.2	771	10	US-09-815-242-7986	Sequence 7986, Ap
20	113	6.2	4315	9	US-09-070-827A-1	Sequence 1, Appl
21	112	6.1	753	9	US-09-738-626-2865	Sequence 2865, Ap
22	112	6.1	3309400	9	US-09-738-626-1	Sequence 1, Appl
23	110	6.0	654	9	US-09-738-626-1499	Sequence 1499, Ap
24	110	6.0	969	9	US-09-938-842A-329	Sequence 329, Ap
25	104	5.7	5793	10	US-09-880-107-2109	Sequence 2109, Ap
26	101	5.5	699	10	US-09-974-300-1012	Sequence 1012, Ap
27	101	5.5	1406	10	US-09-939-980-119	Sequence 119, Ap
28	99	5.4	726	10	US-09-815-242-8504	Sequence 8504, Ap
29	99	5.4	30365	10	US-09-825-414-1	Sequence 1, Appl
30	98.5	5.4	702	10	US-09-815-242-4259	Sequence 4259, Ap
31	97.5	5.3	2313	9	US-09-938-842A-1580	Sequence 1580, Ap
32	97	5.3	801	12	US-10-007-693-50	Sequence 50, Appl
33	97	5.3	3309400	9	US-09-738-626-1	Sequence 1, Appl
34	95.5	5.2	1836	9	US-09-938-842A-651	Sequence 651, Ap
35	95.5	5.2	3198	9	US-10-152-661-601	Sequence 601, Ap
36	95.5	5.2	3198	9	US-09-866-050A-601	Sequence 601, Ap
37	94.5	5.2	2040	9	US-09-738-626-1377	Sequence 1377, Ap
38	94.5	5.2	2340	10	US-09-815-242-7995	Sequence 7995, Ap
39	94	5.1	7596	10	US-09-728-952-1	Sequence 1, Appl
40	93	5.1	1500	9	US-09-938-842A-1281	Sequence 1281, Ap
41	93	5.1	2811	9	US-09-938-842A-2538	Sequence 2538, Ap
42	92.5	5.1	291	10	US-09-878-574-8247	Sequence 8247, Ap
43	92	5.0	2670	9	US-09-927-827-27	Sequence 27, Appl
44	92	5.0	23907	9	US-10-077-130-6	Sequence 6, Appl
45	92	5.0	24120	9	US-10-077-130-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-878-574-695
Sequence 695, Application US/09878574
Patent No. US20020110548A1
GENERAL INFORMATION:
APPLICANT: BYRUM, Joseph R.
APPLICANT: LA ROSA, Thomas J.
TITLE OF INVENTION: Thompson, Michael D.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(15401)B
CURRENT APPLICATION NUMBER: US/09/878,574
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 09/333,535
PRIOR FILING DATE: 1999-06-14
NUMBER OF SEQ ID NOS: 15775
SEQ ID NO 695
LENGTH: 252
TYPE: DNA
ORGANISM: Glycine max
OTHER INFORMATION: Clone ID: LIB3028-047-01-B1-D12
US-09-878-574-695

Alignment Scores:
Pred. No.: 1,936-36
Score: 386.50
Percent Similarity: 93.83%
Best Local Similarity: 90.12%

Lengths:
Matches: 252
Conservative: 73
Mismatches: 4

Query Match: 21.12% Indels: 1
DB: 10 Gaps: 1
US-09-857-613A-28 (1-350) x US-09-878-574-695 (1-252)
QY 66 GlnysgylleallagluPhetyraspGluSerSerglyLeuTrpGluasnIletrpgly 85
DB 9 CAGAAGGAAATGACAGAGTCTACACAGCGTCTGCGATGTGGAGAACATTTGGCGC 68
QY 86 ASPHSMethIshIsglyPhetyraspSerSerpTrhValSerLeuSeraspHsary 105
DB 69 GATCATATGACACAGCGCTTTATATGACCGGATTCACCGCTTCTGTTCTATCATCTCC 128
QY 106 AlaalaGlnIleargMetIleGlnIleuSerLeuArgpheAlaSer--ValSerGluGln 124
DB 129 GCTGCTCAGATCCGAAATGATCCAGAAATCTCTGTTTGCTCTTCTGCTGAGAAC 188
QY 125 ArgSerlystrpProlySerIleValaspValIglyCysGlyIleGlyCysSerArg 144
DB 189 CCTTCTAATGCGCCAAAGATATAGTGTGGGTGCGCATTAAGGGCAGCTCCAGA 248
QY 145 Tyr 145
DB 249 TAC 251
RESULT 2
US-09-294-093B-3472
: Sequence 3472, Application US/09294093B
: Patent No. US20010051335A1
: GENERAL INFORMATION:
: APPLICANT: Laigudi, Raghunath, V.
: APPLICANT: Ito, Laura, Y.
: APPLICANT: Sherman, Bradley, K.
: TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
: FILE REFERENCE: PL-0009 US
: CURRENT APPLICATION NUMBER: US/09/294,093B
: PRIOR FILING DATE: 1999-04-16
: PRIOR APPLICATION NUMBER: 60/082,567
: PRIOR FILING DATE: April 21, 1998
: NUMBER OF SEQ ID NOS: 6207
: SOFTWARE: PERL Program
: SEQ ID NO 3472
: LENGTH: 311
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: misc.feature
: OTHER INFORMATION: Incyte ID No. US20010051335A1 700380165H1
: NAME/KEY: unsure
: LOCATION: 4, 43, 55, 96, 99, 127, 156
: OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-3472
Alignment Scores:
Pred. No.: 4.31e-28 Length: 311
Score: 316.00 Matches: 66
Percent Similarity: 76.47% Conservative: 12
Best Local Similarity: 64.71% Mismatches: 23
Query Match: 17.27% Indels: 2
DB: 10 Gaps: 0
US-09-857-613A-28 (1-350) x US-09-294-093B-3472 (1-311)
QY 245 SerLeuHISpTrpGlnGlnaspLeuLeuLysIleCys-AspAlaTyrTyrLeuPr 264
DB 9 TCGCTAAAGCCCGAGTGAACCTGAGCTCGAAGANAAATATGGCGAANCCTACTACTCC 68
QY 264 AlaATrPySerSerTrhAspPyrValLysLeuLeuGlnSerLeuSerLeuGlnAspI 284
DB 69 GGACTGGGCTCAGCTAGCATATGTNANATGGCAAGTACAGTCTCTCGAGATAT 128
QY 284 elysSerGlnsPTrpSerArgPheValAlaProPheTrpProAlaValIleArgSerAl 304

DB 129 CACGACAGCTGACGTGCGAGAACGTCNCCCGCTTTGGCCCGCGTGTATAAATCAGC 188
QY 304 aPheTrpTrpLysGlyLeuSerSerLeuSerSerglyGlnLysThrIleLysGlyAl 324
DB 189 GCTAACATGGAAGGCTTACCTCTGCTGACGACCGCATGGAGACATAGAGCCG 248
QY 324 aLeuAlaMetProLeuMetIleGluGlyTyrLysLysAspLeuIleLysPheAlaIle 344
DB 249 GATGCGATGCGCGTAAATGATCAGGCGTAAAGAGGCGCTCATCA-TTACCATCAT 307
QY 344 eThr 345
DB 308 CACC 311
RESULT 3
US-09-938-842A-1815
: Sequence 1815, Application US/09938842A
: Patent No. US20020160378A1
: GENERAL INFORMATION:
: APPLICANT: Harper, Jeff
: APPLICANT: Kreps, Joel
: APPLICANT: Wang, Xun
: APPLICANT: Zhu, Tong
: TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINI
: FILE REFERENCE: SAME, AND METHODS OF USE
: CURRENT APPLICATION NUMBER: US/09/938,842A
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: US 60/227,866
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: US 60/264,647
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/300,111
: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 5379
: SEQ ID NO 1815
: LENGTH: 1080
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
US-09-938-842A-1815
Alignment Scores:
Pred. No.: 8.89e-17 Length: 1080
Score: 226.00 Matches: 88
Percent Similarity: 42.21% Conservative: 61
Best Local Similarity: 24.93% Mismatches: 150
Query Match: 12.35% Indels: 54
DB: 9 Gaps: 13
US-09-857-613A-28 (1-350) x US-09-938-842A-1815 (1-1080)
QY 31 ValGlyProArgSerTrpAlaProIleArgAlaSerAlaIleSerSerGluArgGlyGlu 50
DB 70 CTAGGTCCACAGAACGAAAGGCAACGAGCTGTGATCTCTCCGC-----GGCTCA 123
QY 51 lleValleGlu-----GlnLysPro 57
DB 124 ATCTCCGCGAAGAAAGTCAAGACACTATTAACCACTAGTGTCTTCTCCGCAACCA 183
QY 58 LysLys--AspAspLysLysLysLeuGlnLysGlyIleAlaGluPheTyrAspGluSer 76
DB 184 AAAGAGATGAGATCAGCCAGAAAGTACTGCTGCGACAGCTTACAGATGTGTGC 243
QY 77 SerGlyLeuTrpGluAsnIleTrpGlyAspHISMetHIS-----Gly 91
DB 244 ACTGATATGTACAGAGTGGGATGGGCAATCTTTCATTTCTCTCCATGCTCCCTGGA 303
QY 92 PheTyrAspSerSerTrhValSerLeuSeraspHsaryAlaAlaGlnIleArgMet 111
DB 304 AATTCGACAAAGACGCCCA-----AGA 327
QY 112 IleGlnLysSerLeuArgPheAlaSerValSerGluGluArgSerLysTrpProLysSer 131


```

Db 328 ATCCAGAGAAATG-----GCCGTCGATCTCATCAAACTGAACCGGACAAAG 378
QY 132 ILevalAspValGlyCysGlyIleGlySerSerArgTyrLeuAlaLysPheGly 151
Db 379 ATTCTTGACCTGCTGGCGGCGGTGGGCGCATGAGACCATTCCGCCATTCCAG 428
QY 152 AlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAsnAlaLeuAla 171
Db 439 GCCCAAGTCACTGATCACTATCAACGAGTACCAAGTGAACGAGCAAGCTTCACAAC 498
QY 172 AlaAlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAlaLeuGlnPro 191
Db 499 AAGAAAGCTGACCTGATTCTCTCTCAACGCTGTTGCTGTAACCTTTTAAAGATGCCG 558
QY 192 PheSerAspGlyGlnPheAspLeuValTyrPheMetGluSerGlyGlnHisMetProAsp 211
Db 559 TTGATGAAACACGCTTTCGAGGAGCTTACTGATCAACCTACGCTGCTCCCTAAG 618
QY 212 LysAlaLysPheValGlyGlnLeuAlaArgValAlaAlaProGlyAlaIleIleIle 231
Db 619 CTCGAAGAGTATACCTCGGAGATCTTCAGAGTATGAACGAGATCTTCTTCGCTCC 678
QY 232 ValThrTyrPcysHisArgAspLeuGlyProAspGlnGlnSerLeuHisProTyrPgluGln 251
Db 679 TACGAATGCGTCAACCTGCAAAATACAGAGCATGACGACGAGAACAC-----AAG 729
QY 252 AspLeuLeuLysLysIleCysAspAlaTyrTyrLeuProAlaTyrPcysSerThrSerAsp 271
Db 730 GACGATATTAAGGATTCGAGAGAGAGACGACGCTTCTGCTGACTAAGACATCGCTGAT 789
QY 272 TyrValLysLeuLeuGlnSerLeuSerLeuGlnAspIleLysSerGluAspTyrSerArg 291
Db 790 ATAGCCGTCGACGGCAAGAAAGTTGGGTTGAGGTAGTGAAGGAAGATTGGCTAAA 849
QY 292 PheValAlaLysProPheTyrProAlaValIle-----ArgSerAlaPheThrTyrLys 308
Db 850 CCACCGCTTAACCGTGGTGAACCGGTTAAAGATGGAGATGGCTTAT---TGAGA 906
QY 309 Gly-----LeuSerSerLeuLeuSerSerGlyGlnThrIleLysGlyAlaLeuAla 326
Db 907 AACCATGTTGCTGCTGATCTTCTTCTGCTATTGGGCTTCTCTTAAGGAGACTGTTGAT 966
QY 327 MetProLeuMetIle-----GluGlyTyrLysAspLeu 338
Db 967 GTTCATAGATGTTGTTTAAAGTCTGATATTGATTTGACGAGAGTGGTGAAGATCGAATC 1026
QY 339 IleLys---PheAlaIleIleThrCysArgLysProGlu 350
Db 1027 TTCTCTCGATGATGATTCTCTGTAGAAACCAAG 1065

RESULT 4
US-09-779-144A-6
: Sequence 6, Application US/09779144A
: Patent No. US20020148006A1
: GENERAL INFORMATION:
: APPLICANT: Nes, David W.
: TITLE OF INVENTION: Transgenic Plants with Modified Sterol Compositions
: FILE REFERENCE: 11899.0198.DVDS01 MOBT:198
: CURRENT APPLICATION NUMBER: US/09/779,144A
: PRIOR APPLICATION NUMBER: US 09/106,926
: PRIOR FILING DATE: 1998-06-29
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 6
: LENGTH: 1497
: TYPE: DNA
: ORGANISM: Zea mays
US-09-779-144A-6

Alignment Scores: 1,47e-16 Length: 1497
Pred. No.: 226.00 Matches: 81
Score:

```

```

Percent Similarity: 38.77% Conservative: 45
Best Local Similarity: 24.92% Mismatches: 121
Query Match: 12.35% Indels: 78
DB: 10 Caps: 10

US-09-857-613a-28 (1-350) x US-09-779-144A-6 (1-1497)

QY 69 ILevalAspValGlyCysGlyIleGlySerSerArgTyrLeuAlaLys 88
Db 216 GTTAAATAATACATGATGATCTTCCACATGCTTCTATGATGATGGTGAATCTTC 275
QY 89 HisHisGlyPheTyrAspSerAspSerThrValSerLeuSerAspHisArgAlaAlaGln 108
Db 276 CACTTCTCTCAACATGAGATGAGATCCCTTACGTAAAGCATCAACGACGACATGACAT 335
QY 109 IleArgMetIleGlnGlnSerLeuArgPheAlaSerValSerGlnGlnArgSerLysTyr 128
Db 336 TTCTCTGCTGCAACCTGCTTGTG-----AACCA 365
QY 129 ProLysSerIleValAspValGlyCysGlyIleGlySerSerArgTyrLeuAlaLys 148
Db 366 GGAATGAAGGTTTATGATGCTGCTGCTGATAGTGGACCATGAGAAATTCGA--- 422
QY 149 LysPheGlyAlaThrSerVal---GlyIleThrLeuSerProValGlnAlaGlnArgAla 167
Db 423 AGATTAGCTCAACTTCAGTACCGGATTCGATTAACAACGATACCATCAACGAGGGA 482
QY 168 AsnAlaLeuAlaAlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAla 187
Db 483 AAGAGCTCAACCGTTTACGAGAAATGATGCAACATGATTTTTCACACGCGACTTC 542
QY 188 LeuGlnGlnProPheSerAspGlyGlnPheAspLeuValTyrPheMetGluSerGlyGlu 207
Db 543 ATGAGATGCCGCTTCGATGACAAACACTTTTATGCTGTTTACGCAATGAGCAACATGT 602
QY 208 HisMetProAspLysAlaLysPheValGlyGlnLeuAlaArgValAlaAlaProGlyAla 227
Db 603 CATGACCTGATCCAGTGTGCTTACAGAGAGATATATCGTGTGTAACCGCGCAG 662
QY 228 IleIleIleIleValIleThrPcys-----His 236
Db 663 TCGTTTCCCGTACGAGTGGTGGCTTACGATCACTATGATCTCAACATGCAACCCAC 722
QY 237 Arg-----AspLeuGly-----ProAspGlnGlnSerLeuHis 247
Db 723 AAAAGATCAAGATGAATGATGACCTTGGCATGCGCCAGATATCAGAAAGCACTGG 782
QY 248 Pro-----TyrGlnGlnAspLeu 254
Db 783 CAATGCTCCGCGCAGTAAAGACCGCGGTTTGAGTTGTTGGATATGAGATCTTCT 842
QY 255 LysLysIleCysAspAlaTyrTyrLeuProAlaTyrPcysSer----- 268
Db 843 GAAATTTCTCCCTTGCCTTGTGCTTCCCTTGGATCCCAACCGATTCCTCCGATAGC 902
QY 269 -----ThrSerAspTyrValLysLeuLeuGlnSer 278
Db 903 TTCCGTTGACCTCTGTCGAGCATGATTAACCCCAATGATCAAGGCCCTGAGATAC 962
QY 279 LeuSerLeuGlnAspIleLysSerGlnLysPyrSerArgPheValAlaProPheTyrPro 298
Db 963 GTTGCTCTGCTCGCAGGAGGAGTGAGGCTCTACTTTCCTGAG----- 1010
QY 299 AlaValIleArgSerAlaPheThrTyrLysGlyLeuSerSerLeuLeuSerSerGlyGln 318
Db 1011 -----AAGCTGCAAGAGGCTGCTGAGCGCGCAAG 1043
QY 319 LysThrIleLysGlyAlaLeuAlaMetProLeuMetIleGluGlyTyrLysAspLeu 338
Db 1044 AAGAGATC-----TTACGCCCAATGATCTTTTCTTTTGTTCGGAACCTCTT 1091
QY 339 IleLysPheAlaIle 343
DB: 10 Caps: 10

```

Db 1092 CTGGAATGAGCTCTT 1106

RESULT 5

US-09-938-842A-1856

Sequence 1856, Application US/09938842A

Patent No. US20020160378A1

GENERAL INFORMATION:

APPLICANT: Harper, Jeff

APPLICANT: Krepes, Joel

APPLICANT: Wang, Xun

APPLICANT: Zhu, Tong

TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

TITLE OF INVENTION: SAME, AND METHODS OF USE

FILE REFERENCE: SCRIPI300-3

CURRENT APPLICATION NUMBER: US/09/938,842A

CURRENT FILING DATE: 2001-08-24

PRIOR APPLICATION NUMBER: US 60/227,866

PRIOR FILING DATE: 2000-08-24

PRIOR APPLICATION NUMBER: US 60/264,647

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/300,111

PRIOR FILING DATE: 2001-06-22

NUMBER OF SEQ ID NOS: 5379

SEQ ID NO 1856

LENGTH: 1086

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-09-938-842A-1856

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
3	35e-15	212.50	42.54%	11.61%	1086	80	54	144	37	11

US-09-857-613a-28 (1-350) x US-09-938-842A-1856 (1-1086)

OY 55 GlnLysProLysLys---AspAspLysLysLysLeuGlnLysGlyLeuAlaGluPheTyr 73

Db 175 CGCGCGTCAAAAGAAATCGAAGACCGCGAAGAGTTCCACATTCGTCGACACATCTTAC 234

OY 74 AspGluSerSerGlyLeuTrpGluAsnLleTrpGlyAspHisMetHisGlyPheTyr 93

Db 235 AATCTCGTCACCGACATATTCAGATGGGATGGGACAAATCTCCAC-----TTCTCA 288

OY 94 AspSerAspSerThrValSerLeuSerAsp-----HisArgAlaAlaGlnLle 109

Db 289 CCATCAATCCCGGAAATCTTCACAAAGACGCCGCCCTCCACGAGATGCCGTA 348

OY 110 ArgMetLleGlnLysLeuSerLeuArgPheAlaSerValSerGluGluArgSerLysTrpPro 129

Db 349 GATCTGATCCAA-----GTCAACCTGGT 372

OY 130 LysSerLleValAspValGlyCysGlyLleGlyLysSerArgTyrLleuAlaLysLys 149

Db 373 CAAAGATCTGATGACGTCGATCGGTGTCGCGGTCGATCGACGAGATTCATCTCAC 432

OY 150 PheGlyAlaThrSerValGlyLleThrLeuSerProValGlnAlaGlnArgAlaAsnAla 169

Db 433 TCCGAGCTAACCTGATCGGATTCACATTAACGAGTATGGAACAGCTCGCTC 492

OY 170 LeuAlaAlaLleGlnLysLeuAlaAspLysValSerPheGlnValAlaAspAlaLleGln 189

Db 493 CACAATTAAGAAAGCTGCTCGACGCGCTTCGAGGTCGTGTGTAACCTCTCCAG 552

OY 190 GlnProPheSerAspCysGlnPheAspLeuValTrpSerMetGluSerGlyGlnHisMet 209

Db 553 ATGCGCTGATGACAAAGACTTTCGACGCGGTATTCATTCAGACGCTGTCACGCG 612

OY 210 ProAspLysAlaLysPheValGlyLleLeuAlaArgValAlaAlaLapProGlyAlaLleTle 229

Db 613 CCGAAGCTGAGAGAGTACCGACAGATCTACAGGCTGTTGAACCCGGATCTATGTAT 672

OY 230 IleLleValThrTrpCysHisArgAspLeuGlyProAspGluGlnSerLleHisProTrp 249

Db 673 GTGTCTGACAGAGTGGCTTACAGCGAACAATTTAAGCCGAGATGACGAAC----- 726

OY 250 GluGlnAspLeuLeuLysLysLleCysAspAlaTyrTyrLeuProAlaTrpCysSerThr 269

Db 727 ---GTGAGCTTAATCCAAAGGATTTGAGAGAGCGGATCGTTACCGAGCTTACGCTTAC 783

OY 270 SerAspPyrValLysLeuLeuGlnSerLeuSerLeuGlnLysPheLysSerGluAspTrp 289

Db 784 GTGATATTAATCTGAGACGCTAAAGGTTGGCTTTGATGATGAGAGAGAGAACATCTG 843

OY 290 SerArgPheValAla---ProPheTrpProAlaValLle-----ArgSerAlaPheThr 306

Db 844 GCGAGTCCACCGGCTGAGCGGCTGAGTGTGACACGTTAAGATGGGTAGGCTTCTAT--- 900

OY 307 TrpLysGly-----LeuSerSerLeuLeuSerSerGlyGlnLysThrLleLysGlyAla 324

Db 901 TGGAGCAATCACAATTGCTGATTTGTACAGATTGTGACGCGTGTGAGTCTCTTAAGCAACT 960

OY 325 LeuAlaMetProLeuMetLle-----GluGlyTyrLysLys 336

Db 961 GTTGATGTTATGATGAGATGTTGTTTAAAGCTGCTGATTTATTTGACCAAGAGAGTGAAACC 1020

OY 337 AspLeuLleLys---PheAlaLleLleThrCysArgLysProGlu 350

Db 1021 GAATATTTCTCTCCGATCATGATGATCTCTGCGAAGAACCGGAG 1065

RESULT 6

US-09-801-368-95

Sequence 95, Application US/09801368

Patent No. US20020128250A1

GENERAL INFORMATION:

APPLICANT: Busby, Robert

APPLICANT: Call, Brian

APPLICANT: Hecht, Peter

APPLICANT: Holzman, Doug

APPLICANT: Madden, Kevin

APPLICANT: Maxon, Mary

APPLICANT: Milne, Todd

APPLICANT: No. US20020128250A1man, Thea

APPLICANT: Royer, John

APPLICANT: Salama, Sofie

APPLICANT: Sherman, Amir

APPLICANT: Silva, Jeff

APPLICANT: Summers, Eric

TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fu

FILE REFERENCE: 109272.147

CURRENT APPLICATION NUMBER: US/09/801,368

CURRENT FILING DATE: 2001-03-07

PRIOR APPLICATION NUMBER: US 09/487,558

PRIOR FILING DATE: 2000-01-19

PRIOR APPLICATION NUMBER: US 60/160,587

PRIOR FILING DATE: 1999-10-20

NUMBER OF SEQ ID NOS: 440

SOFTWARE: PatentIn version 3.0

SEQ ID NO 95

LENGTH: 1152

TYPE: DNA

ORGANISM: Saccharomyces cerevisiae

US-09-801-368-95

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
4	09e-14	203.50	37.99%	11.12%	1152	78	58	149	73	12

US-09-857-613a-28 (1-350) x US-09-801-368-95 (1-1152)

[illegible]

```

RESULT 7
US-09-779-144A-1
; Sequence 1, Application US/09779144A
; Patent No. US20020148006A1
; GENERAL INFORMATION:
; APPLICANT: Nes, David W.
; TITLE OF INVENTION: Transgenic Plants with Modified Sterol Compositions
; FILE REFERENCE: 11899, 0198, 0VUS01, MOBT:198
; CURRENT APPLICATION NUMBER: US/09/779,144A
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: US 09/106,926
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1320
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-779-144A-1

```

Alignment Scores:	
Pred. No.:	5.05e-14
Score:	203.50
Percent Similarity:	37.998
Best Local Similarity:	21.798
Query Match:	11.128
DB:	10
	Gaps:
	12

QY	42	SerAlaAlaSerSerGluArgGlyGluIleValleuGluIuInLysProLysAspasp	61
	
Db	177	AACTCTCCCAAAAGGACGGCTTCGAAAGTACTTGGAATTGGATGGTAGACCGAT	236
QY	62	LysLysLysLeuGlnLysGlyIleAlaGlnPheTyrAspGluSerGlyLeuTrpGlu	81
	: :
Db	237	AAAGATGCCAAGACGCTCGTCTT---GAAGATTATATGAAACCCACACATTCCTACTAT	293
QY	82	AsnIle-----TrpGlyAspHisMetHisGly---Phe	92
		:::	:::
Db	294	AACCTCGTACAGATTTCATGAAATATGGTTGGGGTTCCTTCATTCATTCCACACATTT	353
QY	93	TyrAspSerAspSer-----ThrValSerLeuSerAspHisArgAlaAlaGlnIle	109
		:::	:::
Db	354	TATTAAGGTAGAGTTTCGCTGCCCTGGATRGCAAGACATGAATATTATTTAGGCTTCAAG	413
QY	110	ArgMetIleGlnLysSerLeuArgPheAlaSerValSerGluIuArgSerLysTrpPro	129
Db	414	GCTGGTATTCCAAGAGCG-----	431
QY	130	LysSerIleValAspValGlyCysGlyIleGlyGlySerSerArgTyrLeuAlaLysLys	149
		:::::	:::::
Db	432	GATTAGGTTCTCACCTTGTTGGTGTGGGGGCCAGCAAGAGAGATTTCCAGATT	491
QY	150	PheGlyValAlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAsnAla	169
		:::	:::
Db	492	ACCGGTTTGAAGCTCACTCGGTCTTAACAATTAACGATTACCAAAATTCGCAAGCAAAATAT	551
QY	170	LeuAlaAlaAlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAlaAspAlaLeuGln	189
		:::	:::
Db	552	TACGCTAAAAAATACAAATTGTAGGTGACCAAAATGGAAGCTTGTAAAGGTGATTTTCATGAA	611
QY	190	GlnProPheSerAspGlyLysIlePheAspLeuValTTrpSerMetGluSerGlyLysIleMet	209
		:::	:::
Db	612	ATGAGATTTCGAAGAAACACTTTCGACAAAGTTTATGCAATTGAGAGCCACATGTACCTCT	671
QY	210	ProAspLysAlaLysPheValGlyGluLeuAlaArgValAlaAlaProGlyValAlaIleIle	229
		:::	:::
Db	672	CCAAATTTGAAGGTGTATACAGCGAAATCTACAGAGTTTGAACCGGGGTGATCACTTT	731
QY	230	IleIleValThrTrpCysHisArgAspLeuGlyProAspGluGlnSerLeuHisProTrp	249
		:::::	:::::
Db	732	GCTGTTTACGAATGGTATAGCTATTAATATGACAAACAAATCTCGAATCT-----	785

```

Db      354  TATTAAGGTGAGAGTTTCGTCGCTCGCGATAGCAGACATGAACTATTATTAGCTTACGAG  413
QY      110  ArgMetIleGlnGlnSerLeuArgPheAlaSerValSerGluGluArgSerLysTrpPro  129
      |||||
Db      414  GCTGTATTCTAAAGAGC-----  431
QY      130  LysSerIleValAspValGlyCysGlyIleGlyLysSerArgTrpLeuAlaLys  149
      |||||
Db      432  GATTTACTTCTCGACGTTGGTGGTGGTGGGGGCCACAGACAGATGTCGAATTT  491
QY      150  PheGlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaArgAlaAsnAla  169
      |||||
Db      492  ACCGGTTGTAACGTCATCGCTCTAAACATATACCATTTACCAATTTGCCAGCGAAATAT  551
QY      170  LeuAlaIleAlaGlnGlyIleuAlaAspLysValSerPheGlnValAlaAspAlaLeuGln  189
      |||||
Db      552  TACGCTAAAAAAATACATTTGAGTACGCCAATGAGCTTTTAAAGGGTGATTTCTATATAA  611
QY      190  GlnProPheSerAspGlyLysIlePheAspLeuValTrpSerMetGlnSerGlyGlnHisMet  209
      |||||
Db      612  ATGCAATTTTCGAGAAANAACCTTTCGACAAAGTTTGTGCAATTGAGGCCACATGTCAGCGT  671
QY      210  ProAspLysIleLysPheValGlyGlnLeuAlaArgValAlaAlaIleProGlyAlaIleIle  229
      |||||
Db      672  CCAAAATTAGAAAGGTGTATACACCGCAAACTCTACAAAGTTTGAACCGGGTGTACCTTT  731
QY      230  IleIleValIleThrTrpCysHisArgAspLeuGlyProAspGlnGlnSerLeuHisProTrp  249
      |||||
Db      732  GCTGTTTACCAATGGCAATGACATGATTAATATATGACGAAACCAATTCCTGAACAT-----  785
      |||||
      -Leu 263

```

Db	786	-----AGAAAGATC-----			GCATTGAATTCACTAGCTGATG	824
Qy	264	prolatrtpcysserthnserasptlyvalylsleu	leuglnserleuserleuglnasp			283
Db	825	CCAAAGATGTTCCATGCTGCAGCTGGCTAGGA	AAACATTGAAGAACTGTGGTTTCGAAGTC			884
Qy	284	llelusergiuasp-----	trpserargp	hevala		294
Db	885	CTCGTTAGCGAAGACCTGGCGGACATGATGA	ATCCCTTGCTATTCACCATTA	CTACT		944
Qy	295	prophetpprovalval-----	lleargser	alaph		305
Db	945	GCGACATGAGACGTACGTTCAAAACTTA	CTAATTTGGCCACATTTTCAGAAC	CTTTCAC		1004
Qy	306	thrtiplysglyleuser	ser			
Db	1005	TTGGGATGACAAATTATCTACAGCAATG	TTACTGTAATGAGAAAATTA	GTGCTAG	CCCCCA	1064
Db		-----	leuser			315
Db		-----	leuser			332

Dy 1065 GAAAGTTCCAGGAACCTTACTGCTGTCAAGAAAATGCTGCAGTGGCTTACTTCGCCGT 1124

Dy 333 GLYTYRILYSLSPLEULELYSPHEALALEILETHCYARGLYSPROLU 350

Dy 1125 GGTAAGTCACAAGTTATTCACTCCAATGAHCGTTTTGCTGCATFAGAACCCACAGA 1178

RESULT 9
MS-09-779-1AAA-3

```

US 09/779,144A
; Sequence 3, Application us/0979144A
; Patent No. US20020148006A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Nes, David W.
; TITLE OF INVENTION: Transgenic Plants with Modified Sterol Compositions
; FILE REFERENCE: 11899, 0198, DVSU501 MOBT:198
; CURRENT APPLICATION NUMBER: US/09/779,144A
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: US 09/106,926
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
;
; SEQ ID NO 3

```

```

: LENGTH: 1420
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
US-09-779-144A-3

Alignment Scores:
Pred. No.: 3.7e-13 Length: 1420
Score: 196.50 Matches: 81
Percent Similarity: 42.86% Conservative: 54
Best Local Similarity: 25.71% Mismatches: 143
Query Match: 10.74% Indels: 38
DB: 10 Gaps: 11

US-09-857-613a-28 (1-350) x US-09-779-144A-3 (1-1420)

QY 55 GlnLysProLysLys---AspAspLysLysLysLeuGlnLysLysLysLysLysLys 73
DB 217 CCCCCGCAAAAGAAATCGAAACCCCGAGAAAGTTCCAGACTTCGCGACATCTCTAC 276
QY 74 AspGluSerSerGlyLeuTrpGluAsnIleTrpGlyAspHisMetHisGlyPheTyr 93
DB 277 AATCTCGTCACGACATATACAGAGTGGGAGGACAAATCCTTCCAC-----TTCTCA 330
QY 94 AspSerAspSerThrValSerLeuSerAsp-----HisArgAlaAlaGlnIle 109
DB 331 CCATCAATCCCGGAAATCTCACAAGACGCCAGCCCTCCACGACAGAGATGGCGGTA 390
QY 110 ArgMetIleGlnGluSerLeuArgPheAlaSerValSerGluLysTrpPro 129
DB 391 GATCTGATCCAA-----GTCAACCTGGT 414
QY 130 LysSerIleValAspValGlyCysGlyIleGlySerSerArgTyrLeuAlaLys 149
DB 415 CAAAGATCTTCAAGTCGATCGATCGCTGCGCCATCGACGACGATTCGATCTCAC 474
QY 150 PheGlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAsnAla 169
DB 475 TCGCAGACAGCTA-GTCCGATTTCAATAAACGATACAGTGCACAGAGCTGCTCTC 533
QY 170 LeuAlaAlaAlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAlaLeuGln 189
DB 534 CACAAATGAAGAAAGCTGCTCGACCGCTTTCGAGGCTGCTGTCGTAACCTCCCGAG 593
QY 190 GlnProPheSerAspGlyGlnPheAspLeuValTrpSerMetGluSerGlyLysMet 209
DB 594 ATGCGGTCGATGACACAGAGTTTCGACGAGCTTATTCATCGAAGCCAGCTGATCCG 653
QY 210 ProAspLysAlaLysPheValGlyLeuAlaArgValAlaAlaProGlyAlaIleIle 229
DB 654 CCGAAGCTGGAAGAGTACCCAGAGATCTACAGGCTTGAACCCCGATCTATGTAT 713
QY 230 IleIleValIleTrpCysHisArgAspLeuGlyProAspGluGlnSerLeuHisProTrp 249
DB 714 GTGTCGTCAGAGTGGGTTACGACGAGAAATTAAGCGCGAGAGATGACGACAC----- 767
QY 250 GluGlnAspLeuLeuLysLysIleCysAspAlaTyrTyrLeuProAlaTrpCysSerThr 269
DB 768 ---GTGAGGTAATCCCAAGAGTTGAGAGCGCATGCTTACACAGCGCTTGGGCTTAC 824
QY 270 SerAspTyrValLysLeuLeuGlnSerLeuSerLeuGlnAspIleLysSerGluAspTrp 289
DB 825 GTGATATAGTGAACCGCTTAAAGAGTGGGTTGAGATAGTGAAGAGAGATCTG 884
QY 290 SerArgPheValAla---ProPheTrpProAlaValIle-----ArgSerAlaPheThr 306
DB 885 GCGAGTCCACCGGCTGACCGCTGGGAGCTAGCTTAAGATGGGATGGCTTCTTAT--- 941
QY 307 TrpLysGly-----LeuSerSerLeuLeuSerSerGlyGlnLysThrIleLysGlyAla 324
DB 942 TGGAGGATCAATCATTTGTCAGATTTTTCAGCGGTTGAGTGTCTCCATAAGAACT 1001
QY 325 LeuAlaMetProLeuMetIle-----GluGlyTyrLysLys 336

```

```

DB 1002 GTTGATGCTTCATGAGATGTTGTTAAGACGTCGATTTGTGACGAGAGGATGAACC 1061
QY 337 AspLeuIleLys---PheAlaIleIleThrCysArgLysProGlu 350
DB 1062 GAATATATCTCTCCCATGTCATATGATTTCTGCGAAGAACCGGAG 1106

RESULT 10
US-09-805-681-1
: Sequence 1, Application US/09805681
: Patent No. US20020143163A1
: GENERAL INFORMATION:
: APPLICANT: INNOTE, MASAYORI
: APPLICANT: PHADTARE, SANGITA
: APPLICANT: YAMANAKA, KUNITOSHI
: APPLICANT: KATO, IKUNOSHI
: TITLE OF INVENTION: ANTIBACTERIAL ACTIVITY OF 4,5 DIHYDROXY-2-CYCLOPENTAN-1-ONE
: TITLE OF INVENTION: (DHCP) AND CLONING A GENE CONFERRING DHCP RESISTANCE IN
: FILE REFERENCE: 1137-P-00
: CURRENT APPLICATION NUMBER: US/09/805,681
: PRIOR FILING DATE: 2001-03-14
: PRIOR APPLICATION NUMBER: 60/228,727
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: Patent Ver. 2.1
: SEQ ID NO 1
: LENGTH: 3900
: TYPE: DNA
: ORGANISM: Escherichia coli
US-09-805-681-1

Alignment Scores:
Pred. No.: 8e-08 Length: 3900
Score: 156.50 Matches: 69
Percent Similarity: 41.45% Conservative: 45
Best Local Similarity: 25.09% Mismatches: 105
Query Match: 8.55% Indels: 56
DB: 10 Gaps: 14

US-09-857-613a-28 (1-350) x US-09-805-681-1 (1-3900)

QY 71 GluPheTyrAspGluSerSerGlyLeuTrpGluAsnIleTrpGlyAspHisMetHis 90
DB 2100 GAGCATTTACGATTTGGATGATGATGTTGACCGGATCTTGATCCCTCATGCAATAT 2159
QY 91 Gly-----PheTyrAspSerAspSerThrValSerLeuSerAspHisArgAlaAla 107
DB 2160 TCCTCGCTTACTGGAAGATCCGAT-----AATCTGGAATCTGCCAGCAGCGG 2210
QY 108 GlnIleArgMetIleGlnGluSerLeuArgPheAlaSerValSerGluLysTrpSerLys 127
DB 2211 AAGCTCAAAATGATTTTGAAATAATTCAGTTA-----AAA 2246
QY 128 TrpProLysSerIleValAspValGlyCysGlyIleGlyGlySerSerArgTyrLeuAla 147
DB 2247 CCAGGATCGCGGTATCGGATGATTTGCTGCGCTGGCGGAGCTGGACACATGCGCA 2306
QY 148 LysLysPheGlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAla 167
DB 2307 TCTAATTTATGACGTGAAGCGTGGGCGTCCACATTTCTGCCGGAACGCAAAAAATGCT 2366
QY 168 AsnAlaLeuAlaAlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAla 187
DB 2367 CAG-----GAACGCTGTAAGGCTGGATGTCACCAAT--- 2399
QY 188 LeuGlnGlnProPheSerAsp-----GlyGlnPheAspLeuValTrpSerMetGluSer 205
DB 2400 TTGCTGCAAGATTTATCGGACCTGACAGCAGATTTGATGCTATTTCTGCTGGGATG 2459
QY 206 GlyLysHisMetProAspLys-----AlaLysPheValGlyGluLeuAlaArgValAla 223
DB 2460 TTCGAGCAGCGTGAGACCGCAAAATAATAGATATTTTGGGCTGGATGCTAATTTG 2519

```

QY	184	ValAlaSpAlaLeuGlnGlnProPheSerSpGlyGlnPheAspLeuValTrpSerMet	203
Db	163	TGTGCGGATTCACCTAAAAAACATACCCCTAGATACATCTGATGTATATATTCCTCGCT	222
QY	204	GluserGlyGluHisMetProAspIysAlaIysPheValGlyGluLeuAlaArgValAla	223
Db	223	GACACATGTTCATACATCATCAAGATTAACCAACCACCATATTCAGATCATTTTACAAATGGCTTG	282
QY	224	AlaProGlyAlaIleIleIleValIleThrProCysHisArgAspLeuGlyPro	241
Db	283	AAGCGTGGAGGTACACTTATTTACTGATTCATGCAAAAGTGAAGAAAGTGTCA	336
RESULT 12			
	US-10-260-877-53		
	Sequence 53, Application US/10260877		
	Publication No. US20030021813A1		
	GENERAL INFORMATION:		
	APPLICANT: Abbott Laboratories		
	APPLICANT: Chovan, Linda E.		
	APPLICANT: Hessler, Paul E.		
	APPLICANT: Reich, Karl A.		
	TITLE OF INVENTION: ESSENTIAL BACTERIA GENES AND GENOME		
	TITLE OF INVENTION: SCANNING IN HAEMOPHILUS INFLUENZAE FOR THE IDENTIFICATION OF		
	TITLE OF INVENTION: 'ESSENTIAL GENES'		
	FILE REFERENCE: 6565 US P1		
	CURRENT APPLICATION NUMBER: US/10/260,877		
	CURRENT FILING DATE: 2002-09-30		
	PRIOR APPLICATION NUMBER: US/09/649,145		
	PRIOR FILING DATE: 2000-08-25		
	NUMBER OF SEQ. ID NOS: 137		
	SOFTWARE: FastSeq for Windows Version 4.0		
	SEQ ID NO 53		
	LENGTH: 753		
	TYPE: DNA		
	ORGANISM: H. influenzae		
	FEATURE:		
	NAME/KEY: CDS		
	LOCATION: (1)...(753)		
	OTHER INFORMATION: HI-0095		
	US-10-260-877-53		
Alignment Scores:			
	Pred. No.: 4.02e-07	Length: 753	
	Score: 141.00	Matches: 47	
	Percent Similarity: 44.68%	Conservative: 37	
	Best Local Similarity: 25.00%	Mismatches: 74	
	Query Match: 7.70%	Indels: 30	
	DB: 9	Gaps: 5	
US-09-857-613A-28 (1-350) x US-10-260-877-53 (1-753)			
QY	130	LysSerIleValAspValGlyCysGlyIleGlyIleGlySerSerArgTrpLeuAlaIys	149
Db	118	AAAAAAGTGTGGAGGTTCCTGCTTAATATGGGACGACATGCAATATGGATTCGCAAAACAA	177
QY	150	PheGlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAspAla	169
Db	178	TTTGCTGTCTATTTGAAGCTGTGATTTAATGATCAATGAAAAATGCGTTAGCAAAAGCACAAGCA	237
QY	170	LeuAlaAlaGlnGlyLeuAlaAspIysValSerPheGlnAlaAlaAspAlaLeuGln	189
Db	238	AATATTGAAGCAAAATGGCTTCAGGAAAAAATTCATGTACAGCGTGGCAATGCGATGAAAG	297
QY	190	GlnProPheSerAspGlyGlnPheAspLeuValTrpSerMetGluSerGlyGluHisMet	209
Db	298	TTGCTTTCGAGATGAAGATTTCATATTCATCATCAATGAAGCAGCATGCTCAATGTATA	357
QY	210	Pro-----AspIysAlaIysPheValGlyGluLeuAlaArgValAlaAlaProGlyAla	227
Db	358	CCCGTGAAGCAAGAAAAAGCATTTGCGAGATATATTTCGAGTGTAAATACCATGCT	417
QY	228	IleIleIleIleValIleThrTrpCysHisArgAspLeuGlyProAspGlnGlnSerLeuHis	247

Db 418 TTATTGCTT-----ACTCAGCATGTTATGCTGGGGAATGATCATCAACTATT--- 468
Qy 248 ProtrpGlInaSpLeuLysLysLleCysAspAlaTyrTyrLeuProAlaTrpPys 267
Db 469 -----CTAATAAATATATGCCAAGCGATT----- 492
Qy 268 SerThrSerAspTyrValLysLeuLeuGlnSerLeuSerLeuGlnAspLleLysSerGlu 287
Db 493 -----AACGTGACCTGTCAGCGCCATTACCAAGAT 522
Qy 288 AspTrpSerArgPheValAlaProPheTrpProAlaValIleArgSerAlaPheThrTrp 307
Db 523 GGATGGAAAGCGCAT-----TTCCAAGAAAGTGTTTGAATAATGTTGATACCTTCC 573
Qy 308 LysGlyLeuSerSerLeuLeuSer 315
Db 574 TCTGTGAGATGACATTACTTCTCC 597

RESULT 13
US-09-845-248-1
: Sequence 1, Application US/09845248
: Patent No. US2002042106A1
: GENERAL INFORMATION:
: APPLICANT: Nampoothiri, Madhavan
: TITLE OF INVENTION: Nucleotide Sequences Which Code for the CMA Gene
: FILE REFERENCE: 032301 WD 1161
: CURRENT APPLICATION NUMBER: US/09/845,248
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 1801
: TYPE: DNA
: ORGANISM: Corynebacterium glutamicum
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (248)..(1600)
: OTHER INFORMATION:
US-09-845-248-1

Alignment Scores:
Pred. No.: 2,02e-06 Length: 1801
Score: 140.00 Matches: 83
Percent Similarity: 36.06% Conservative: 45
Best Local Similarity: 23.38% Mismatches: 120
Query Match: 7.65% Indels: 107
Gaps: 15

US-09-857-613A-28 (1-350) x US-09-845-248-1 (1-1801)

Qy 53 LeuGlnGlnLysPro-----LysLysAspAsp 61
Db 674 ATGGAACAAGCCCTGCATGCGCTAAAGCATCATCAGCGGCTAGCATCCAGCTCG 733
Qy 62 LysLysLysLeuGlnLysGlyIleAlaGluPheTyrAspGluSerSerGlyLeuTrpGlu 81
Db 734 AATATCCCGGACAGAAAGCAATACCTACTACCTAGCAGCGGCGCATGAGTTCTACTCC 793
Qy 82 AsnIleTrpGlyAspHisMetHis-----GlyPheTyrAspSerAspSerThrVal 99
Db 794 CTGTTTATAGATGATTCCTACGTACCTATACCTCGCGCTATTATCAACGCCAGATCA--- 850
Qy 100 SerLeuSerAspHisArgAlaIleAlaGlnIleArgMetIleGlnLysLeuArgPheAla 119
Db 851 AGTTTGAAGAACCCCAAGAAACAAATACCGCTCATCTTTGAAATAACTGGCTCG--- 907
Qy 120 SerValSerGluGlnArgSerLysSerIleValAspValGlyCysGlyIle 139
Db 908 -----AAAGAAAGCGCATCCCTCTAGACGTGGATCGCGTTGG 946
Qy 140 GlyIysSerArgTyrTyrLeuAlaLysPheGlyAlaThrSerValGlyIleThrLeu 159
Db 947 GGAGCGATGATGTCCTTAC---GCCGCCAACAACAGCGTGTGAAGAACCATCGGATTAACGCTG 1003

Qy 160 SerProValGlnAlaGlnArgAlaAsnAlaLeuAlaIleAlaGlnGlyLeuAlaAspLys 179
Db 1004 TCTGAACAGCATATGATGTTGGGTCAACAGATCAACAGCCAAAGTTTGAAGACCTC 1063
Qy 180 ValSerPheGlnValAlaAspAlaLeuGlnInProPheSerAspGlyGlnPheAspLeu 199
Db 1064 GCGAATATTCCTTCATGATGATTACCGCATTTCCA-----GAACCTGATTCGATGCG 1117
Qy 200 ValTrpSerMetGluSerGlyGlnHisMet-----ProAsp----- 211
Db 1118 ATCTCAGCAATCGCATCATGATGACACATCGGTGTAACAACATTCCTCCGACTTTGAA 1177
Qy 212 -----LysAlaLysPheValGlyGluLeuAla----- 220
Db 1178 TTGCTCAGCAGCAACATCTCAACAGCGGCGAGTGAAGTCAACCAACATCACCATTACCA 1237
Qy 221 -----ArgValAlaAlaProGlyAlaIleIleIleIleValIleThrTrpCysHisArgAsp 238
Db 1238 GACAACCGCCCGCCGACGCGAGTGCATTATT-----GATCGCTAC 1279
Qy 239 LeuGlyProAsp----- 242
Db 1280 ATTTTCCCGACGCGTGAACCTACTGCTGCGCCTGATCAAGCATGCAAGCAAC 1339
Qy 243 -----GlnGlnSerLeuHisPro 248
Db 1340 CGTTTGAAGTCTGCACGACAGAAACCTCCGCTTGATTACCAACGACCCCTGACGCG 1399
Qy 249 TrpGlnGlnAspLeuLeuLysLysLleCysAspAlaTyrTyrLeuProAla----- 265
Db 1400 TGGTGGAAACCTCAAGAAATTTGGAGAGACAGTTGAACCTGCGCGTGAACCCACT 1459
Qy 266 -----TrpCysSerThrSerAspTyrVal 273
Db 1460 GCACGACTCTTGGCCCTGACATGCGAGGTTGCGAATGGGATTGCCCAACACATCGTC 1519
Qy 274 LysLeuLeuGlnSerLeuSerLeuGlnAspLleLysSerGluAspTrpSerArgPheVal 293
Db 1520 CAGCTGCACCAAGTACTGCTGTGA---CTCGATGACAGGGAAGTCGCGGAGAA 1573
Qy 294 AlaPro-----PheTrpProAlaValIleArg--SerAlaPheThrTrpLysGlyLeu 311
Db 1574 GTTCCTGAAGAAAGTGGTGCATCATCTAAGAAACAATGTTCTTTTAAG----- 1625
Qy 311 rSerLeuLeuSerSerGlyGlnLysThrIleLysGlyAlaLeu 325
Db 1626 -----GTTCCCTAGGGGCGAGAGTTAATTGTAGTAACTC 1661

RESULT 14
US-09-738-626-636
: Sequence 636, Application US/09738626
: Publication No. US20020197605A1
: GENERAL INFORMATION:
: APPLICANT: NAKAGAWA, SATOSHI
: APPLICANT: NAKAGAWA, SATOSHI
: APPLICANT: MIZOGUCHI, HIROSHI
: APPLICANT: ANDO, SEIKO
: APPLICANT: HAYASHI, MIKIRO
: APPLICANT: OCHIAI, KEIKO
: APPLICANT: YOKOI, HARUHIKO
: APPLICANT: TATEISHI, NAOKO
: APPLICANT: SENOH, AKIHIRO
: APPLICANT: IKEDA, MASAKI
: APPLICANT: OZAKI, AKIO
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
: FILE REFERENCE: 249-125
: CURRENT APPLICATION NUMBER: US/09/738,626
: PRIOR FILING DATE: 2000-12-18
: PRIOR APPLICATION NUMBER: JP 99/377484
: PRIOR FILING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: JP 00/159162
: PRIOR FILING DATE: 2000-04-07
: PRIOR APPLICATION NUMBER: JP 00/280988

PRIOR FILING DATE: 2000-08-03
 NUMBER OF SEQ ID NOS: 7059
 SOFTWARE: PatentIn ver. 3.0
 SEQ ID NO 636
 LENGTH: 1353
 TYPE: DNA
 ORGANISM: Corynebacterium glutamicum
 US-09-738-626-636

Alignment Scores: 1.7e-06 Length: 1353
 Pred. No.: 139.00 Matches: 63
 Score: 37.108 Conservative: 42
 Percent Similarity: 22.26% Mismatches: 122
 Best Local Similarity: 7.60% Indels: 56
 Query Match: 9 Gaps: 12

US-09-857-613a-28 (1-350) x US-09-738-626-636 (1-1353)

53 LeuGlucInLysPro-----LysLysAspAsp 61
 427 ATGGACACACCCCTGATGCGCTAAAGCACATGATCAAGCGCTAGCATCCAGCAGCTCG 486
 62 LysLysLysLeuGlnLysGlyIleAlaGluPheTyrAspGluSerSerGlyLeuTrrpGlu 81
 487 AAATCCCGCAGACAGAAAGCCATTAGCTACACATACGACGCGGCAATGATTTCTACTCC 546
 82 AsnIleTrrpGlyAspHisMetHisHis-----GlyPheTyrAspSerAspSerThrVal 99
 547 CTGTTTATGATGATTCATGATGACCTACCTCGCGCTATATCCAAAGCCAGATCA---- 603
 100 SerLeuSerAspHisArgAlaAlaAlaGlnIleArgMetIleGluSerLeuArgPheAla 119
 604 AATTGGAGAGAACCCCAAGAAACAAATACCGCTCATCTTTGAAAAAAGCTGCTG--- 660
 120 SerValSerGluGlnArgSerLysTrrpLysSerIleValAspValGlyCysGlyIle 139
 661 -----AAGAAGCGGATCGCTCTCTAGACGCGGATCGCTGG 699
 140 GlyGlySerSerArgTrrpLeuAlaLysLysPheGlyValArgSerValGlyIleThrLeu 159
 700 GAGAGCATGTGCTCGCTAC-----GCCGCCAAACACGCTGTGAAAGCCAGGATTCACCTG 756
 160 SerProValAlaAlaGlnArgAlaAsnAlaLeuAlaAlaGlnGlyLeuAlaAspLys 179
 757 TCTGAACAGCAATGATGATGCGGTCAAGCCAGATCAAGCCCAAGCTTTGGAAGACCTC 816
 180 ValSerPheGlnValAlaAlaAspAlaLeuGlnInProPheSerAspGlyGlnPheAspLeu 199
 817 GCGGAATTCGCTTCATGATGATTCACCGCATGCTTCCA-----GAAACTGATTCGATCGC 870
 200 ValTrrpSerMetGluSerGlyGlnHisMet-----ProAspLysAlaLys 214
 871 AATCTGACGCAATGCGCATCTTGAACACATCGGTGACACACTATCCGACTACTTTGAA 930
 215 PheValGlyLeuLeuAlaArgValAlaAlaAlaProGlyAlaIleIleIleValTrrp 234
 931 TTGCTACGACGCAAACTCAAAACA-----GCGGACTGATGCTC 969
 235 CysHisArgAspLeuGlyProAspGluGlnSerLeuHisProTrrpGluAspLeu 254
 970 AACCAACGACATCACTACCCGACACACCGCCGCCAC-----GCA 1011
 255 LysLysIleCysAspAlaTyrTrrpLeuPro-----AlaTrrpCysSerThrSerAspTyr 272
 1012 GGTGATTTATTTGATCGCTACATTTTCCGACAGGTGAACTCACTGCTGACACCTG 288
 273 ValLysLeuLeuGlnSerLeuSerLeuGlnAspIleLysSerGluAsp----- 288
 1072 ATCAAGCAATGACAGACAGAGGTTTGCAGAGTGTGACAGAAAGAACTCCGCTTGTAT 1131
 289 -----TrrpSerArgPheValAlaAlaProPheTrrpProAlaValIle 301

Db 1132 TACCAAGCACCTGTCAGCGCTGTGCGAANAACCTCAAGAAATTTGGAGAGACAGTT 1191
 Oy 302 ArgSerAla 304
 Db 1192 GAACATCGCC 1200

RESULT 15
 US-09-294-093B-5022
 Sequence 5022, Application US/09294093B
 Patent No. US20010051335A1
 GENERAL INFORMATION:
 APPLICANT: Ito, Laura, Y.
 APPLICANT: Sherman, Bradley, K.
 TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
 FILE REFERENCE: PL-0009 US
 CURRENT APPLICATION NUMBER: US/09/294, 093B
 CURRENT FILING DATE: 1999-04-16
 PRIOR APPLICATION NUMBER: 60/082,567
 PRIOR FILING DATE: April 21, 1998
 NUMBER OF SEQ ID NOS: 6207
 SOFTWARE: PERL Program
 SEQ ID NO 5022
 LENGTH: 290
 TYPE: DNA
 ORGANISM: Zea mays
 FEATURE:
 NAME/KEY: misc-feature
 OTHER INFORMATION: Incyte ID No. US20010051335A1 700355594H1
 US-09-294-093B-5022

Alignment Scores: 3e-06 Length: 290
 Pred. No.: 128.00 Matches: 28
 Score: 55.428 Conservative: 18
 Percent Similarity: 33.73% Mismatches: 35
 Best Local Similarity: 6.99% Indels: 2
 Query Match: 10 Gaps: 1

US-09-857-613a-28 (1-350) x US-09-294-093B-5022 (1-290)

Oy 130 LysSerIleValAspValGlyCysGlyIleGlyLysSerArgTrrpLeuAlaLysLys 149
 20 CAGAAAGTACTGATGATGCGATGTGGAATTTGAGCGCGGACTTTTACATGCTCAAAAC 79
 Oy 150 PheGlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAsnAla 169
 80 TATGATGTTCATGTTCTTGTTGTTGATCTTTCCGTCAACATGTT-----TCATTTGCA 133
 Oy 170 LeuAlaAlaAlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAlaLeuGln 189
 134 ATTGAACGTGCGCATTTGCGACGCAAGTCTCTGTTGAATTCGAATTCCTGATTCACACCA 193
 Oy 190 GlnProPheSerAspGlyGlnPheAspLeuValTrrpSerMetGluSerGlyGlnHisMet 209
 194 AAGGATTACCAAGAAATATTGATTTGACGTGATCAGCCGCTGACACACATCCCTTACATA 253
 Oy 210 ProAspLys 212
 254 CAAGCAAG 262

Search completed: March 19, 2003, 09:06:49
 Job time : 100 secs.

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 19, 2003, 06:53:46 ; Search time 1289 Seconds
(without alignments) 4397.533 Million cell updates/sec

Title: US-09-857-613a-28

Perfect score: 1820
Sequence: 1 MATVVRIRPISCIHIFHFRS.....IEGYKKDLIKFALITCRKPE 350

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O/cgpn2.1/USPTO.spool/US09857613/rnat_12032003_100036_6027/app-query.fasta.1.519
-DB=EST -QPM=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09857613 @CGN.1.1456.@rnat_12032003_100036_6027 -MCP=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-MAIN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6 -FCGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estbda:*
2: em_esthum:*
3: em_estinu:*
4: em_estnu:*
5: em_estov:*
6: em_estro:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_yrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1136.5	62.1	1304	11	AY109799	AY109799 Zea mays
2	1029	56.2	599	14	BQ299297	BQ299297 sa045f1l
3	1024	56.0	669	10	BE331080	BE331080 so95a05.Y
4	980	53.6	567	14	BQ273543	BQ273543 sa02e1l
5	975	53.3	570	14	BQ283302	BQ283302 sa06c02
6	971.5	53.1	778	12	BG127606	BG127606 EST47316
7	964	52.7	554	13	BM527813	BM527813 sa16b10
8	957	52.3	556	13	BI469324	BI469324 sa110d1l
9	946	51.7	704	13	BI469281	BI469281 sa109h01
10	931	50.9	831	12	BJ288530	BJ288530 BJ288530
11	922	50.4	664	13	BG320098	BG320098 Zm03.01d1
12	899	49.1	664	13	BJ463750	BJ463750 BJ463750
13	884	48.3	515	14	BQ273302	BQ273302 sa02e12
14	878	48.0	532	10	AM596072	AM596072 s197e09.Y
15	869	47.5	628	9	AI487374	AI487374 EST245696
16	855	46.7	581	13	BI321569	BI321569 sa112h01
17	844	46.1	509	12	BG511373	BG511373 sa115a05
18	840	45.9	484	13	BI321583	BI321583 sa115a05
19	836	45.7	708	14	BQ801716	BQ801716 WHE2817.G
20	824	45.0	503	12	BG790250	BG790250 sa067a12
21	810	44.3	649	13	BI953869	BI953869 HYSME001
22	807	44.1	707	14	BG416854	BG416854 HYSME001
23	794	43.4	691	14	BQ116842	BQ116842 EST602418
24	788	43.1	540	10	BE331113	BE331113 so95906.Y
25	777	42.5	517	12	BG046436	BG046436 sa04a02
26	775	42.3	587	10	BE340311	BE340311 EST344382
27	765	42.0	627	12	BG522228	BG522228 19-89 Ste
28	763	41.7	708	12	BG416864	BG416864 HYSME001
29	759	41.5	439	13	BE804492	BE804492 s197e09.Y
30	754	41.2	444	13	BM094790	BM094790 sa121002
31	734.5	40.1	653	10	AM774108	AM774108 EST333338
32	728	39.8	553	12	BE631792	BE631792 NF00803D
33	720	39.3	587	10	BE591261	BE591261 WHE16505-1
34	710	38.8	661	13	BJ447900	BJ447900 BJ447900
35	707	38.6	733	12	BF051091	BF051091 EST436266
36	685.5	37.5	533	9	AI489112	AI489112 EST247451
37	670	36.6	517	10	BE607043	BE607043 WHE0915_H
38	665	36.3	484	10	AM617931	AM617931 EST314005
39	662.5	36.2	739	12	BG523936	BG523936 37-16 Ste
40	660	36.1	630	10	AV945401	AV945401 AV945401
41	658	36.0	491	9	AI485917	AI485917 EST244238
42	657	35.9	531	10	BE426891	BE426891 WHE0333-C
43	656	35.8	452	10	AM099184	AM099184 s035h03.Y
44	650	35.5	623	12	BG447567	BG447567 NF004H01S
45	642	35.1	449	12	BG725508	BG725508 sa038c12

ALIGNMENTS

RESULT 1
AY109799
LOCUS AY109799
DEFINITION Zea mays CL364_1 mRNA sequence.
ACCESSION AY109799
VERSION AY109799.1 GI:21213657
KEYWORDS
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; PAC
Clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 1304)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitlitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.

```
|||||.....F---L-----S-----C-----P-----A-----I-----D-----S-----P-----A-----L-----E-----U-----G-----I-----N-----190
```

source 1. .599

```

/organism="Glycine max"
/db.xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl054-7197"
/clone_lib="Gm-cl054"
/tissue_type="Leaf, 3 week old, greenhouse grown"
/lab_host="DH10B"
/notes="Vector: Bluescript II SK-, Site.1: EcoRI; Site.2:
XhoI; The Harosoy NIL was constructed and seed was
provided by Dr. J. Specht, University of Nebraska
(Shoemaker and Specht, 1995). The cDNA library was
constructed from mRNA isolated from leaf tissue at various
developmental stages of 3 week old greenhouse grown
plants. Complementary DNA was synthesized from mRNA using
a primer consisting of a poly(dT) sequence with a XhoI
restriction site and a 3' anchor. EcoRI adapters were
ligated to the blunt-ended cDNA fragments followed by
XhoI digestion. The cDNA fragments were directionally
cloned into the EcoRI-XhoI restriction site of the
bluescript vector. The ligated cDNA fragments were
transformed into DH10B host cells (GibcoBRL). This
library was constructed in cooperation with Dr. Paul
Keim's laboratory at Northern Arizona University."

BASE COUNT      139 a      144 c      165 g      151 t
ORIGIN

Alignment Scores:
Pred. No.:      6,27e-96      Length:      599
Score:          1029.00      Matches:      199
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    56,23%      Indels:      0
DB:             14          Gaps:         0

US-09-857-613a-28 (1-350) x BQ299297 (1-599)

QY      6 ArgIleProThrIleSerCysIleHisIleHisThrPheArgSerGlnSerProArgThr 25
        |||
        3 AGATATCCCAACATTCATCATGCATCCACATCCACGCTCCGTCCTCCCAATCCCTCCACCT 62
QY      26 PheAlaArgIleArgValGlyProArgSerTrpAlaProIleArgAlaSerAlaIleSer 45
        |||
        63 TTCGCCAGATCCGGGCTGGAGACCAGCTGCTGGCTTCCTATTCGGGCTCCGACGAGCAGC 122
QY      46 SerGluArgGlyGlnIleValIleGluGlnIleValProLysProLysAspLysLysIleu 65
        |||
        123 TCGGAGAGAGGGAGATGATGATTGAGCAGAACCCGAAAGAGATGCAAGAAAGAGCTG 182
QY      66 GlnLysGlyIleAlaGlnPheTyrAspGluSerSerGlyLeuTrpGlnuSnIleTrpGly 85
        |||
        183 CAGAAAGGAAATCGACAGATTTCACGACGCTCTTCGCTTATGGGCAACATTTGGGCGC 242
QY      86 AspHisMetHisIleGlyPheTyrAspSerAspSerThrValSerLeuSerAspHisArg 105
        |||
        243 GACCAATCACCACATCGCTTATGCTGATTCACATCTTCCTTCGATCATCTCT 302
QY      106 AlaAlaGlnIleArgMetIleGlnIleSerLeuArgPheAlaSerValSerGluArg 125
        |||
        303 GCTGCTCAGATCCGATGATCCAAAGATCTCTTCGCTTCCTGCTTCTGAGAGCCT 362
QY      126 SerLysTrpProLysSerIleValAspValGlyCysGlyIleGlyIleSerSerArgTyr 145
        |||
        363 AGTAAATGGCCCAAGAGATGATGATGATGGGTGGCATGGGCGACGCTCTGATATC 422
QY      146 LeuAlaLysLysPheGlyValAlaThrSerValGlyIleThrLeuSerProValGlnIleArg 165
        |||
        423 CTGGCCCAAGAAATTTGGAGCAACAGCTAGCATCTCTAGTCCCTGTTCAACTCA 482
QY      166 ArgAlaAsnAlaLeuAlaAlaGlnIleLeuAlaAspLysValSerPheGlnValAla 185
        |||
        483 AGAGCAAAATGCTCTGCTGCTCAAGGATGGCTGATAGGTTCTCTTCAGGTTGCT 542
QY      186 AspAlaLeuGlnGlnProPheSerAspGlyGlnPheAspLeuValTrpSerMetGlu 204
        |||

```

```

DB      543 GACGCTACAGCAACCATTCCTGACGGCCAGTTGATCTGTGTGTCATGAG 599
RESULT 3
LOCUS   BE331080
DEFINITION BE331080 669 bp mRNA linear EST 04-DEC-2001
          sc95a05.y1 Gm-cl041 glycine max cDNA clone GENOME SYSTEMS CLONE ID:
          Gm-cl041-1065 5' similar to TR:09ZSK1 09ZSK1 GAMMA-T0COPHEROL
          METHYLTRANSFERASE. ; mRNA sequence.
ACCESSION BE331080
VERSION   BE331080.1
KEYWORDS EST.
SOURCE   BE331080.1 GI:9204856
ORGANISM Glycine max
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
          Glycine.
REFERENCE 1 (bases 1 to 669)
          Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Corveill,V., Rhanna
          ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
          Wylie,T., Underwood,K., Steptoe,M., Rheising,B., Allen,M., Bowers
          ,Y., Peterson,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
          ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
          ,R., Waterston,R. and Wilson,R.
          Public Soybean EST Project
          Unpublished (1999)
          Contact: Shoemaker R/Public Soybean EST Project
          Public Soybean EST Project
          Washington University School of Medicine
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
          Tel: 314 286 1800
          Fax: 314 286 1810
          Email: est@watson.wustl.edu
          This clone is available through: Resgen, Invitrogen Corp. 2130
          South Memorial Parkway Huntsville, AL 35801 For further information
          call: (800)-533-4363 or contact via email: c@resgen.com
          High quality sequence stop: 461.
          Location/Qualifiers
            1.669
              /organism="Glycine max"
              /db.xref="taxon:3847"
              /clone="GENOME SYSTEMS CLONE ID: Gm-cl041-1065"
              /clone_lib="Gm-cl041"
              /tissue_type="Senescing leaves, mature plant, greenhouse
              grown"
              /lab_host="DH10B"
              /note="Vector: pTrp3pac (Pharmacia); Site.1: EcoRI;
              Site.2: HindIII. This library was constructed from mRNA
              isolated from senescing leaf tissue of mature greenhouse
              grown plants of the cultivar Williams. Complementary DNA
              was synthesized from mRNA using a 3' anchored poly(dT)
              primer. EcoRI adapters were ligated to the blunt-ended
              cDNA fragments followed by digestion with EcoRI and
              HindIII. The cDNA fragments were directionally cloned
              into the EcoRI-HindIII restriction site of the pTrp3-pac
              vector. The ligated cDNA fragments were transformed into
              DH10B host cells. This library was constructed by Dr.
              Randy Shoemaker."

BASE COUNT      154 a      156 c      184 g      174 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:      2.37e-95      Length:      669
Score:          1024.00      Matches:      200
Percent Similarity: 97.10%      Conservative: 1
Best Local Similarity: 96.62%      Mismatches: 6
Query Match:    55.96%      Indels:      0
DB:             10          Gaps:         0

US-09-857-613a-28 (1-350) x BE331080 (1-669)

QY      1 MetAlaThrValAlaArgIleProThrIleSerCysIleHisIleHisThrPheArgSer 20
        |||

```


Db 423 ATATATAGTACATGGTGGCAGGAGATCTGGCCCTGACAGACAATCTTACATCA 482

QY 249 TTPGLUGLNASPLEULELysLYSILECYSAPLALATYTYRLEUPROALATPcysSer 268
|||||

Db 483 TGGAGCAGAGCTCTTAACAGACATTTGGCATGATTAATCTCCCTGGCTGTCTCA 542
|||||

QY 269 ThisSerAspTYrValLYSLeuLeu 276
|||||

Db 543 ACTCTGATTAATGTTAGTGTCTC 566
|||||

RESULT 5
B0786302

LOCUS DEFINITION B0786302 570 bp mRNA linear EST 26-JUN-2002
sag66c02.y1 Gm-cl076 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-cl076-4708 5' similar to TR:09XHP9 09XHP9 GAMA-TOCOPHEROL
METHYLTRANSFERASE. [1] mRNA sequence.

ACCESSION B0786302

VERSION B0786302.1 GI:21994774

KEYWORDS

SOURCE soybean.

ORGANISM Glycine max

REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 570)
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corryell, V., Khanna
Wylie, T., Underwood, K., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: ccregen@resgen.com web site:
www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 422.

FEATURES

Source Location/Qualifiers

1..570

/organism="Glycine max"

/db_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl076-4708"

/clone_id="Gm-cl076"

/tissue_type="wounded cotyledons"

/dev_stage="11 day old seedlings"

/lab_host="DH10B"

/note="Vector: pBluescript II SK(+). Site_1: EcoRI; Site_2:
XhoI. The cDNA library was constructed from mRNA isolated
from 11 day old seedlings treated with that were treated
with 2 ugs/ml of a crude glucan elicitor preparation
isolated from the mycelial walls of Phytophthora sojae.
The library was prepared using the Stratagene pBluescript
II SK(+) library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with a blunt-ended cDNA fragments
adaptors were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA fragments were
directionally cloned into the EcoRI-XhoI restriction site
of the pBluescript vector. The ligated cDNA fragments were
transformed into E.coli ElectroMax DH10B host cells. Plant
material was provided by Michael G. Hahn (Complex
Carbohydrate Research Center, University of Georgia) and
the library was constructed by Anu Khanna (Lila Vodkin Lab

BASE COUNT 156 a 123 c 136 g 155 t

ORIGIN

Alignment Scores:

Pred. No.: 2,16e-90

Score: 975.00

Percent Similarity: 99.47%

Best Local Similarity: 98.93%

Query Match: 53.28%

DB: 14

Gaps: 0

US-09-857-613a-28 (1-350) x B0786302 (1-570)

QY 164 ALaGlnArGAlaAsnaLaLeuAlaAlaAlaGlnGlyLeuAlaAspLYSValSerPheGln 183
|||||

Db 1 GCTTAAGACCAATGCTCTGCTGCTGCACGACGATGCTGATTAAGCTTCCCTTCCAG 60
|||||

QY 184 ValAlaAspAlaLeuGlnGlnInProPheSerAspGlyInPheAspLeuValTTPSerMet 203
|||||

Db 61 GTTGCTAGCGCTCTACAGCAGACCATTTCTTGACGGCCAGTTTATCTGGTGGTCCATG 120
|||||

QY 204 GlnSerGlyGlnHisMetProAspLYSAlaLYSPheValGlyLeuLeuAlaArgValAla 223
|||||

Db 121 GAGAGTGGAGACGATATGCTGCACAAGCTAAGTTTGTGGAAGTACCTCGGGTACGA 180
|||||

QY 224 AlaProGlyValAlaLeuLeuLeuLeuValThrTTPCysHisArgAspLeuGlyProAspGlu 243
|||||

Db 181 GCACCAAGGTCCACTATTAATAATGATGACATGCTGCCACAGGATCTGGCCCTGACGAA 240
|||||

QY 244 GlnSerLeuHisProTTPGlnGlnAspLeuLeuLYSILECYSAPLALATYTYRLeu 263
|||||

Db 241 CAATCCTTACATCCATGAGGAGCAGATCTCTTAAGAACAATTTGCGATCAATTAACCTC 300
|||||

QY 264 ProAlaTTPCysSerThrSerAspTYrValLYSLeuLeuGlnSerLeuSerLeuGlnAsp 283
|||||

Db 301 CCGCGCTGGCTGCACACTTCTGATTAATGTTAAGTCTCCCAATCCCTGCACCTCAGGC 360
|||||

QY 284 IleLYSergLYAspTYrSerArgPheValAlaArgPheTTPProAlaValIleArgSer 303
|||||

Db 361 ATCAAGTCCGAAGATTTGGCTGCTTGTCTTCCATTTTGGCCAGCAGATGATACGCTCA 420
|||||

QY 304 AlaPheThrTTPGlyGlyLeuSerSerLeuSerGlyGlnLYSHrIleLYSGly 323
|||||

Db 421 GCCTTCACATGGAAGGCTCTACTTCACTTGCACAGTGGCAAAACGATTAAGAAGA 480
|||||

QY 324 AlaLeuAlaMetProLeuMetIleGlnGlyTYrLYSAspLeuIleLYSPheAlaIle 343
|||||

Db 481 GCTTGGCTATGCCATTTGATGATACAGGATTCAGAAAGATCTAATTAAGTTCCATC 540
|||||

QY 344 IleThrGysArgGlyProGlu 350
|||||

Db 541 ATTACATGTGCAAAACCTGAA 561
|||||

RESULT 6

LOCUS B6127606 778 bp mRNA linear EST 31-JAN-2001

DEFINITION EST473168 tomato shoot/meristem Lycopersicon esculentum cDNA clone

ACCESSION B6127606

VERSION B6127606.1 GI:12627710

KEYWORDS EST.

SOURCE tomato.

ORGANISM Lycopersicon esculentum

REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.

AUTHORS van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Utterback, T.,
Hansen, C., Rönning, C. and Tanksley, S.

TITLE Generation of ESTs from tomato shoot/meristem tissue

JOURNAL Unpublished (2001)

ORIGIN

Alignment Scores: 2.84e-89 Length: 554
 Pred. No.: 964.00 Matches: 183
 Score: 99.468
 Percent Similarity: 99.468
 Best Local Similarity: 99.468
 Query Match: 52.688
 DB: 13 Gaps: 0

US-09-857-613A-28 (1-350) x BM527813 (1-554)

OY 68 GYIIEALGIuphetYAspIuserSerGIyleuTrpGIuAsnIleTrpGIyAspHis 87
 DB 2 GGAATCGAGAGTTTACAGACAGCTTCTGCTTATGAGAACATTTGGCGACAC 61
 OY 88 MethIsHsIGlypHetYAspSerSerThValSerIuserSerAspHisArgAla 107
 DB 62 ATGACCATGGCTTTATGACTCGGATTCACATGTTTGGCTTTCGATCATGCTGCT 121
 OY 108 GlnIleArgMetIleGlnIuserIleuArgPheAlaSerValSerGIuArgSerLys 127
 DB 122 CAGATCCGAATGATCCAGAGCTCTTCTGCTTCTGCTTCTGAGAGCTAGTAA 181
 OY 128 TrpProLysSerIleValAspValGIyScGIyIleGIySerSerArgTrpLeuAla 147
 DB 182 TGGCCCAAGACTATATGATGATGTTGGTGGCATAGCTGACACTATACCTGGCC 241
 OY 148 LysIlypPheGIyAlaTrpSerValGIyIleTrpIuserProValGlnAlaArgAla 167
 DB 242 AAGAAATTTGAGACACAGCTGATGCTCTGATCTGCTGCTGCTGCTGCTGCTGCT 301
 OY 168 AsnAlaLeuAlaAlaIleGlnIleuAlaAspLysValSerPheGlnValAlaAspAla 187
 DB 302 AATGCTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
 OY 188 LeuGlnIleuPheSerSerAspGIyGlnPheAspLeuValTrpSerMetGIuSerGIyGlu 207
 DB 362 CTACAGACACATTTCTGAGCGCCAGCTGATGATGCTGCTGCTGCTGCTGCTGCT 421
 OY 208 HisMetProAspLysAlaLysPheValGIyGIuLeuAlaArgValAlaAlaProGIyAla 227
 DB 422 CATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 481
 OY 228 IleIleIleIleValIleTrpCysHisArgAspLeuGIyProAspGIuGlnSerLeuHis 247
 DB 482 ACTATATATATAGTACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541
 OY 248 ProTrpGIuGln 251
 DB 542 CCATGGGAGCAA 553

RESULT 8

BI469324

LOCUS

DEFINITION sail0011.y1 Gm-cl053 Glycine max cDNA clone GENOME SYSTEMS CLONE

ACCESION ID: Gm-cl053-2854.5' similar to TR:09XIP9 09XIP9 GAMMA-TOCOPHEROL

VERSION

KEYWORDS BI469324.1 GI:15285433

SOURCE

ORGANISM

soybean.

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine

1 (bases 1 to 562)

REFERENCE

AUTHORS

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corryell, V., Rhana,

A., Bolla, B., Matra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,

Wille, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers,

Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schuck,

R., Rlter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann

TITLE

JOURNAL Public Soybean EST Project
 COMMENT Contact: Shoemaker R/Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu

FEATURES
 This clone is available through: Resgen, Invitrogen Corp. 2130
 South Memorial Parkway Huntville, AL 35801 For further information
 call: (800)-533-4363 or contact via email: ccuresgen.com
 High quality sequence stop: 421.

SOURCE

1. 562
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl053-2854"
 /clone_jib="Gm-cl053"
 /tissue_type="Whole seedling, 3 week old, greenhouse
 grown"
 /lab_host="DH10B"
 /note="Vector: pBluescript II SK+; Site.1: EcoRI; Site.2:
 XhoI; The Harosoy NIL was constructed and seed was
 provided by Dr. J. Specht, University of Nebraska
 (Shoemaker and Specht, 1995). The cDNA library was
 constructed from mRNA isolated from whole seedlings of 3
 week old greenhouse grown plants. Complementary DNA was
 synthesized from mRNA using a primer consisting of a
 poly(dT) sequence with a XhoI restriction site and a 3'
 anchor. EcoRI adapters were ligated to the blunt-ended
 cDNA fragments followed by XhoI digestion. The cDNA
 fragments were directionally cloned into the EcoRI-XhoI
 restriction site of the pBluescript vector. The ligated
 cDNA fragments were transformed into DH10B host cells
 (GibcoBRL). This library was constructed in cooperation
 with Dr. Paul Keim's Laboratory at Northern Arizona
 University."

BASE COUNT 130 a 136 c 153 g 143 t

ORIGIN

Alignment Scores:
 Pred. No.: 1.52e-88 Length: 562
 Score: 957.00 Matches: 186
 Percent Similarity: 99.478
 Best Local Similarity: 99.478
 Query Match: 52.308
 DB: 13 Gaps: 0

US-09-857-613A-28 (1-350) x BI469324 (1-562)

OY 11 SerCysIleHsIleHsIleHsThrPheArgSerGlnSerProArgThrPheAlaArgIleArg 30
 DB 2 TCATGACATCCACATCCACAGCTTCGTTCCCAATCCCTCCGACTTTGCCAGATCCGG 61
 OY 31 ValGIyProArgSerTrpAlaProIleArgAlaSerAlaIleSerSerGIuArgGIyGlu 50
 DB 62 GTCGAGCCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
 OY 51 IleValLeuGlnIleuLysProLysLysAspAspLysLysLysLysLysLysLysLys 70
 DB 122 ATAGTATTTGGAGCAAGAGCCGAGAGAGATGACAGAGAGAGAGAGAGAGAGAGATGCA 181
 OY 71 GluPheTrpAspGIuSerSerGIyLeuTrpGIuAsnIleTrpGIyAspHisMetHis 90
 DB 182 GAGTTTACGACGAGCTCTTGGCTATGCGAGAACATTTGGCGGACACATGACCAT 241
 OY 91 GlyPheTrpAspSerSerThValSerLeuSerAspHisArgAlaIleArg 110
 DB 242 GCGTTTATATACCTGATTCACATGCTTTCGCTTCCGATCATGCTGCTGCTGCTGCTGCA 301
 OY 111 MetIleGlnIleuSerLeuArgPheAlaSerValSerGIuLysSerLysTrpProLys 130


```

|||||
Db 302 ATGATCCAGAGCTCTGCTTCCCTTGGCTTCTGAGAGCTAGTAATGCGCCAG 361
|||
Qy 131 SerileValaspValaGlyCysGlyIleGlyGlySerSerArgTyrLeuAlaLysLysSph 150
|||
Db 362 AGTATAGTATGATGTTGGTGGTGGATAGTGGAGCTCTGATGATCTGGCCAGAAATTT 421
|||
Qy 151 GlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnAlaAspAlaLeu 170
|||
Db 422 GGACCAACCGAGTGAAGCATCCTCTGCTCAAGCTCAAGAGCAAAATGCTCTT 481
|||
Qy 171 AlaAlaAlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAlaLeuGlnGln 190
|||
Db 482 GCTGCTGCTCAAGGATGCTGATAGGCTTCTTCCAGTTGCTGACGCTCTACAGCA 541
|||
Qy 191 ProPheSerAspGlyGlnPhe 197
|||
Db 542 CCATTCCTCTGACGGCCAGTTT 562
|||
RESULT 9 556 bp mRNA linear EST 30-NOV-2001
LOCUS B1469281
DEFINITION B1469281.1 GI:15285390
ID: Gm-c1053-3025 5' similar to TR:Q9XIP9 Q9XIP9 GAMMA-TOCOPHEROL
METHYLTRANSFERASE. [1] ; mRNA sequence.
ACCESSION B1469281
VERSION B1469281.1 GI:15285390
KEYWORDS
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 556)
Shoemaker, R., Kelm, P., Vodkin, L., Erpelting, J., Corvett, V., Khanna
, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wylie, T., Underwood, K., Stepcoe, M., Theising, B., Allen, M., Bowers
, K., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Hunttsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: coudresgen.com
High quality sequence stop: 422.
Location/Qualifiers
1..556
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1053-3025"
/tissue_type="Whole seedling, 3 week old, greenhouse
grown"
/lab_host="DH10B"
/note="vector: Bluescript II SK+, Site_1: EcoRI; Site_2:
XhoI; The Harosoy NIL was constructed and seed was
provided by Dr. J. Specht, University of Nebraska
(Shoemaker and Specht, 1993). The cDNA library was
constructed from mRNA isolated from whole seedlings of 3
week old greenhouse grown plants. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with a XhoI restriction site and a 3'
anchor. EcoRI adaptors were ligated to the blunt-ended
cDNA fragments followed by XhoI digestion. The cDNA

```

```

fragments were directionally cloned into the EcoRI-XhoI
restriction site of the plasmid vector. The ligated
cDNA fragments were transformed into DH10B host cells
(GibcoBRL). This library was constructed in cooperation
with Dr. Paul Keim's laboratory at Northern Arizona
University."
BASE COUNT 129 a 135 c 152 g 140 t
ORIGIN
Alignment Scores:
Pred. No.: 2,04e-87 Length: 556
Score: 946.00 Matches: 184
Percent Similarity: 99.46% Conservative: 0
Best Local Similarity: 99.46% Mismatches: 1
Query Match: 51.69% Indels: 0
Gaps: 0
DB: 13
US-09-857-613a-28 (1-350) x B1469281 (1-556)
Qy 11 SerCysIleHisIleHisThrPheArgSerGlnSerProArgThrPheAlaArgIleArg 30
|||
Db 2 TCATGCATCCACATCCACACGTTCCCTCCCAATCCCTCGCACTTCCGCAAGATCCG 61
|||
Qy 31 ValGlyProArgSerTrpAlaProIleArgAlaSerAlaAlaSerSerGluArgGlyGlu 50
|||
Db 62 GTCGACCCAGTGCCTGGCTCTATTCCGGCATCGCCAGCAGCTCGAGAGAGGGAG 121
|||
Qy 51 IleValLeuGlnGlnLysProLysLysAspAspLysLysLysLeuGlnLysGlyIleAla 70
|||
Db 122 ATAGTATTGAGCAGAGAGCCGAGAGAGATGACAGAGAGAGAGAGAGAGAGAGAG 181
|||
Qy 71 GluPheTyrAspGlnSerSerGlyLeuTrpGluAsnIleTrpGlyAspHisMetHis 90
|||
Db 182 GAGTTTACGACGAGCTTCTTCCGCTTATGAGAGAACTTGGCGGACACATGCACTC 241
|||
Qy 91 GlyPheTyrAspSerAspSerThrValSerLeuSerAspHisArgAlaAlaGlnIleArg 110
|||
Db 242 GCTTTTATGATCGGATTCACATGCTTCCGTTGGATCATGCGTCTCAGATCCGA 301
|||
Qy 111 MetIleGlnIleSerLeuArgPheAlaSerValSerGluGluArgSerLysTrpProLys 130
|||
Db 302 ATGATCCAGAGTCTCTTCCCTTCCCTTCTGTTCTGAGAGAGAGAGAGAGAGAGAG 361
|||
Qy 131 SerileValaspValaGlyCysGlyIleGlyGlySerSerArgTyrLeuAlaLysLysSph 150
|||
Db 362 ACTATAGTATGATGTTGGTGGTGGATAGTGGCAGCTTACATCTGCGCCAGAAATTT 421
|||
Qy 151 GlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnAlaAspAlaLeu 170
|||
Db 422 GGAGCAACCGAGTGAAGCATCCTGAGTTCGTTCACAGCTCAAGAGCAAAATGCTCTT 481
|||
Qy 171 AlaAlaAlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAlaLeuGlnGln 190
|||
Db 482 GCTGCTGCTCAAGGATGCTGATAGGCTTCTTCCAGTTGCTGACGCTCTACAGCA 541
|||
Qy 191 ProPheSerAspGly 195
|||
Db 542 CCATTCCTCTGACGGC 556
|||
RESULT 10 704 bp mRNA linear EST 09-APR-2002
LOCUS B1288530
DEFINITION B1288530 Y. Ogihara unpublished cDNA library, Wh. SL Triticum
aestivum cDNA clone whs114p22 5', mRNA sequence.
ACCESSION B1288530
VERSION B1288530.1 GI:20106970
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae
; Triticaceae; Triticum.
1 (bases 1 to 704)
REFERENCE

```


US-09-857-613a-28 (1-350) x BQ2730298 (1-831)

QY 116 LeuArgpheaAlaSerValSerLincLuarGserLysTrpProLysSerLysValAspVal 135
 DB 827 CTTCCGCGGTTCCACGCCCTCATGATCCAGAGACACCAAAAACATATAGTCGATGC 768
 QY 136 GlyCysGlyIleGlyValSerSerArgTyrLeuAlaLysLysPheGlyAlaThrSerVal 155
 DB 767 GGTGGTGGCATGTGGTGTGATCTCAAGTACTGGCCAAAGATATACGAGCGCAGTGCAC 708
 QY 156 GlyIlePheLeuSerProValGlnAlaGlnArgAlaAsnAlaLeuAlaIleGlnGly 175
 DB 707 GGGATCACGTTGAGCCCTGTTCAAGCCGAGAGGAAATGCTCTCGTGGACGCCAGGG 648
 QY 176 LeuAlaAspLysValSerPheGlnValAlaAspAlaLeuGlnInpPheSerAspGly 195
 DB 647 TTGTCCGATCAGGTACTCTGCAAGTGTGATGCTCTGGAGCAACGGTTCTCCGACGG 588
 QY 196 GlnPheAspLeuValTrpSerMetGluSerGlyLysIleMetProAspLysAlaLysPhe 215
 DB 587 CAGTTCGATCTGGTGTGCTCATGAGAGTGGCAGACATGCCGACAAAGAAAGTTT 528
 QY 216 ValGlyLysLeuAlaArgValAlaAlaProGlyAlaIleIleIleValThrTrpCys 235
 DB 527 GTTAGTGAGCTAGACGCGCTGCGGCTCTGAGAGGACAAATATATCTGACATGTCG 468
 QY 236 HisArgAspLeuGlyProAspGluGlnSerLeuHisProTrpGluGlnAspLeuLys 255
 DB 467 CATAGAACCTGGATCCATCCGAAACCTGGCTAAAGCCCGATGAACTGAGCCTCTGAG 408
 QY 256 LysIleCysAspAlaTrpTyrLeuProAlaTrpCysSerThrSerAspTyrValLysLeu 275
 DB 407 AGCATATGCGACGCGCTACTACTCCGGAGCTGTGCTCAGCTTCAAGACTTGAACATT 348
 QY 276 LeuGlnSerLeuSerLeuGlnAspLysLysSerGluAspTrpSerArgPheValAlaPro 295
 DB 347 GCCAAGTACTGCTCTCAGAGATATCAAGACAGTCTGTCGAGACACGTCGCGCCG 288
 QY 296 PheTrpProAlaValIleArgSerAlaPheThrTrpLysGlyLeuSerSerLeuSer 315
 DB 287 TTTTGGCCGCGCTGATTAATATCAGCGCTTACATGGAAGGCTTCACTCTGCTGAG 228
 QY 316 SerGlyGlnLysThrIleLysGlyAlaLeuAlaMetProLeuMetIleGluGlyTyrLys 335
 DB 227 ACCGATGGAAGACGATCAGAGCGGCGATGTCATGCCCTAATGATCAGCGGCTACAG 168
 QY 336 LysAspLeuLysPheAlaIleIlePheCysArgLysPro 349
 DB 167 AAGGGGCTCATCAAAATTCACCATCATCAGCTGTCGCAAGCT 126
 RESULT 12
 BQ463750 664 bp mRNA linear EST 23-MAY-2002
 LOCUS BQ463750 K. Sato unpublished cDNA library, cv. Haruna Nijo
 DEFINITION germination shoots Hordeum vulgare subsp. vulgare cDNA clone
 bags30f10 5', mRNA sequence.
 BQ463750.1 GI:21142257
 EST.
 SOURCE Hordeum vulgare subsp. vulgare.
 ORGANISM Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 ; Triticeae; Hordeum.
 REFERENCE 1 (bases 1 to 664)
 AUTHORS Sato, K., Saito, D. and Takeda, K.
 TITLE Barley EST sequencing project in NIG and Okayama Univ
 JOURNAL Unpublished (2002)
 COMMENT Contact: Tadasu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856

FEATURES
 source
 1..664
 /organism="Hordeum vulgare subsp. vulgare"
 /cultivar="Haruna Nijo"
 /db_xref="taxon:112509"
 /clone_lib="K. Sato unpublished cDNA library, cv. Haruna
 Nijo germination shoots"
 /tissue_type="shoots"
 /dev_stage="germination"

BASE COUNT 158 a 177 c 186 g 143 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,8e-82 Length: 664
 Score: 899.00 Matches: 168
 Percent Similarity: 86.24% Conservative: 20
 Best Local Similarity: 77.06% Mismatches: 30
 Query Match: 49.13% Indels: 0
 DB: 13 Gaps: 0

US-09-857-613a-28 (1-350) x BQ463750 (1-664)

QY 73 TyrAspGluSerSerGlyLeuTrpGluAsnIleTrpGlyAspHisMetHisGlyPhe 92
 DB 4 TACAGCAGTGCCTCCGCGCTGTGGAGACATCTGGCGCAGACATGCACACCGCTTC 63
 QY 93 TyrAspSerAspSerThrValSerLeuSerAspHisArgAlaAlaGlnIleArgMetIle 112
 DB 64 TACGACTGTGGCGAGCGCGCTTCATGTCCGACACCGCGCGCCAGATCCGATGATC 123
 QY 113 GlnGluSerLeuArgPheAlaSerValSerGluGluArgSerLysTrpProLysSerIle 132
 DB 124 GAGAGCGCCCTGCTTCCCGCGCGCCGAGACGCCCAACCAACCAACCAATT 183
 QY 133 ValAspValGlyCysGlyIleGlyLysSerSerArgTyrLeuAlaLysLysPheGlyAla 152
 DB 184 GTTGATGTGGATGCGCAATCGCGGTGTGCTCAAGTACTGCGCAAAATATGAGCA 243
 QY 153 ThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAsnAlaLeuAla 172
 DB 244 CAGTCTCTGGATCACATTCAGCCAGCTGCAAGCGAGAGAGAAATCCCTTAGCGCG 303
 QY 173 AlAGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAlaLeuGlnInpPhe 192
 DB 304 GCACAGGGGTTGGCAGACAAAGGCTTCTTCCAAAGTCTGATGCTGAGACACATT 363
 QY 193 SerAspGluGlnPheAspLeuValTrpSerMetGluSerGlyLysIleMetProAspLys 212
 DB 364 CCTGATGGCAGTTTGATCTTGTGCTGTGAGAGAGTGTGAGCAGCATGCCAACA 423
 QY 213 AlAlasPheValGlyGluLeuAlaArgValAlaAlaProGlyAlaIleIleVal 232
 DB 424 CAGAAGTGTGAGTGCAGCTGACGCGCTCCAGCTCCAGGAGCAACATATCATCTG 483
 QY 233 ThrTrpCysHisArgAspLeuGlyProAspGluGlnSerLeuHisProTrpGluInp 252
 DB 484 ACCGTGTCATAGCAACCTTCGCTATCTGAGCACTCATGMAACCTGACAGCTGAT 543
 QY 253 LeuLeuLysLysIleCysAspAlaTrpTyrLeuProAlaTrpCysSerThrSerAspTyr 272
 DB 544 CTTTGAATAAATTTGTGATCATATACCTCCGCGATTTGCTCGCCCTCGCATTTT 603
 QY 273 ValLysLeuLeuGlnSerLeuSerLeuGlnAspLysIleLysSerGluAspTrpSer 290
 DB 604 GTCAAGATTGCCAGATCATTTCTTGGATATCAAAACGGCTGACTGCT 657
 RESULT 13
 BQ273302 515 bp mRNA linear EST 07-MAY-2002
 LOCUS BQ273302
 DEFINITION saoz3e12.y1 Gm-c1054 Glycine max cDNA clone SOYBEAN CLONE ID:

Gm-c1054-5207 5' similar to TR:Q9XIP9 Q9XIP9 GAMMA-TOCOPHEROL
METHYLTRANSFERASE. [1] , mRNA sequence.

ACCESSION BQ273302
VERSION BQ273302.1 GI:20498372
KEYWORDS EST.
SOURCE soybean.

ORGANISM Glycine max

Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:
Rosidae: eurosids I: Fabales; Fabaceae: Papilionoideae; Phaseoleae:
Glycine.

REFERENCE 1 (bases 1 to 515)

AUTHORS

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna
A., Bolla, B., Marr, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers,
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk,
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
R., Waterston, R. and Wilson, R.
Public Soybean EST Project

TITLE Unpublished (1999)
JOURNAL Contact: Shoemaker R/Public Soybean EST Project
COMMENT Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp., 2130
South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact: cculeresgen.com web site:
www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 421.

FEATURES
source Location/Qualifiers

1..515

/organism="Glycine max"

/db_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-c1054-5207"

/clone_1lb="Gm-c1054"

/tissue_type="leaf, 3 week old, greenhouse grown"

/lab_host="DH10B"

/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; The Harosoy NIL was constructed and seed was
provided by Dr. J. Specht, University of Nebraska
(Shoemaker and Specht, 1995). The cDNA library was
constructed from mRNA isolated from leaf tissue at various
developmental stages of 3 week old greenhouse grown
plants. Complementary DNA was synthesized from mRNA using
a primer consisting of a poly(dT) sequence with a XhoI
restriction site and a 3' anchor. EcoRI adapters were
ligated to the blunt-ended cDNA fragments followed by
XhoI digestion. The cDNA fragments were directionally
cloned into the EcoRI-XhoI restriction site of the
pBluescript vector. The ligated cDNA fragments were
transformed into DH10B host cells (GibcoBRL). This
library was constructed in cooperation with Dr. Paul
Keim's laboratory at Northern Arizona University."

BASE COUNT 124 a 113 c 135 g 143 t
ORIGIN

Alignment Scores:

Pred. NO: 4.56e-81

Score: 884.00

Percent Similarity: 99.42%

Best Local Similarity: 99.42%

Query Match: 48.31%

DB: 14

US-09-857-613a-28 (1-350) x BQ273302 (1-515)

OY 89 HSHHSGlypHeTyrASpSerAspSerThValSerLeuSerAspHisArgAlaIaGln 108
|||||
Db 3 CACCATGGCTTTTAAAGATCGGATTCACACTGTTGCGTTGAGATCATTGCTGCTCAG 62

OY 109 ILeArgMetIleGlnGlnSerLeuArgPheAlaSerValSerGlnGlnArgSerLysTrp 128
|||||
Db 63 ATCCGATATGATCCAGAGTCCTTCGCTTCCTCTGTTCTTAGACGCTACTAATATG 122
OY 129 ProLysSerIleValAspValGlyCysGlyIleGlyIleGlySerSerArgTyrLeuAlaLys 148
|||||
Db 123 CCCAAGAGTATAGTTGATGTTGGGTGGCATAGTGCGCAGCTGATACCTGCGCCAG 182
OY 149 LysPheGlyValAlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAsn 168
|||||
Db 183 AATTGGAGCAACGAGCTAGCATCCTGAGCTGCTGCTCAAGCAAGCAAAAT 242
OY 169 AlAlaValAlaAlaGlnGlnLeuAlaAspLysValSerPheGlnValAlaAspAlaLys 188
|||||
Db 243 GCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
OY 189 GlnGlnProPheSerAspGlyGlnPheAspLeuValTyrSerMetGlnSerGlyLysHis 208
|||||
Db 303 CAGCAACCATTTCTCTGACGCGCAGTTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362
OY 209 MetProAspLysAlaLysPheValGlyGlnLeuAlaArgValAlaAlaProGlyAlaIle 228
|||||
Db 363 ATGCTGCAAGCAAGCTAAGTTGTTGGAGAGTTAGCTCGGGTAGCACACACAGTGCCTACT 422
OY 229 ILeIleIleValThrTrpCysHisArgAspLeuGlyProAspGlnGlnSerLeuHisPro 248
|||||
Db 423 ATAAATATATGTAACATGATGCGCCACAGAGATCTTGGCCCTGACGACACATCTTACATCCA 482
OY 249 TrpGlnGlnAspLeuLeuLysLysIleCysAsp 259
|||||
Db 483 TGGAGCAAGATCTCTTAAGAAGATTTCGAT 515

RESULT 14

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 532)

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna

A., Bolla, B., Marr, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,

Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers,

Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk,

R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann

R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

CONTACT: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: Resgen, Invitrogen Corp., 2130

South Memorial Parkway Huntville, AL 35801 For further information

call: (800)-533-4363 or contact via email: cculeresgen.com

Insert Length: 1413 Std Error: 0.00

High quality sequence stop: 429.

FEATURES

source Location/Qualifiers

1..532

/organism="Glycine max"

/db_xref="taxon:3847"

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 19, 2003, 05:12:15 ; Search time 2336 Seconds

(without alignments)
4360,438 Million cell updates/sec

Title: US-09-857-613A-28

Perfect score: 1830

Sequence: 1 MATVVRIPITSCIHIFRFS.....IEGYKKDLIKFAITCRKPE 350

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n model -DEV-xmlh
-O-/cgn2.1/USPTO/US09857613/runat_12032003.100036.6013/app_query.fasta_1.519
-DB-GenBml -QFMT-fastlap -SUFFIX-rige -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=45
-UNITS-bits -STRAT-1 -END-1 -MATRIX-biosum62 -TRANS-human4.0.cdi -List=45
-LOCALIGN=200 -THR.SCORER=100 -THR.MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09857613@cgn.1.1.1687-etunal_12032003.100036.6013 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_NMAP -LARGEQUERY -NEG.SCORER=0 -WAIT -LONSDLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6 -FCGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenBml:.*
1: gb_da:.*
2: gb_hlg:.*
3: gb_in:.*
4: gb_om:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_sts:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vl:.*
15: em_da:.*
16: em_fun:.*
17: em_hum:.*
18: em_in:.*
19: em_jmu:.*
20: em_om:.*
21: em_or:.*
22: em_ov:.*
23: em_pat:.*
24: em_ph:.*
25: em_pl:.*
26: em_ro:.*
27: em_sts:.*
28: em_un:.*

29: em_vl:.*
30: em_hlg_hum:.*
31: em_hlg_inv:.*
32: em_hlg_other:.*
33: em_hlg_mus:.*
34: em_hlg_pln:.*
35: em_hlg_rod:.*
36: em_hlg_mam:.*
37: em_hlg_vrt:.*
38: em_sy:.*
39: em_hlgo_hum:.*
40: em_hlgo_mus:.*
41: em_hlgo_other:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1159	63.3	1369	AF213481	AF213481 perilla f
2	1154.5	63.1	1047	AY090280	AY090280 Arabidops
3	1154.5	63.1	1303	AY049258	AY049258 Arabidops
4	1151.5	62.9	1350	AX417728	AX417728 Sequence
5	1151.5	62.9	1350	AF104220	AF104220 Arabidops
6	1147.5	62.7	1296	AY087138	AY087138 Arabidops
7	981.5	53.6	11835	AC006193	AC006193 Arabidops
8	758	41.4	125422	AP003994	AP003994 Arabidops
9	659	36.0	343550	AP003587	AP003587 Nostoc sp
10	647.5	35.4	134199	SYCSLRF	SYCSLRF Nostoc sp
11	579	31.6	108553	AF527809	AF527809 Sorghum b
12	416.5	22.8	298750	AP005375	AP005375 Thermosyn
13	392.5	21.4	334520	AP003588	AP003588 Nostoc sp
14	392	21.4	930	AX073663	AX073663 Sequence
15	392	21.4	957	AX073657	AX073657 Sequence
16	392	21.4	957	AX417730	AX417730 Sequence
17	392	21.4	974	AX073659	AX073659 Sequence
18	392	21.4	145709	D90914	D90914 Synchocyst
19	294	16.1	25681	SAE414559	SAE414559 Saccharot
20	287	15.7	63734	AF127374	AF127374 Streptom
21	282.5	14.8	90445	AF040570	AF040570 Amycolato
22	270	14.7	125561	NC966	NC966 Neurospor
23	268.5	14.3	2122	A60301	A60301 Sequence 1
24	261	14.3	2122	AR144762	AR144762 Sequence
25	261	14.3	2122	NT071107	NT071107 Nicotiana t
26	247	13.5	1276	AB032524	AB032524 Streptom
27	244.5	13.4	5998	AX250261	AX250261 Sequence
28	242	13.2	30000	AX250261	AX250261 Sequence
29	241	13.2	1392	TAU06754	TAU06754 Trifolium ae
30	237	13.0	1410	E07846	E07846 DNA sequenc
31	234.5	12.8	349498	AP003002	AP003002 Mesorhizo
32	232.5	12.7	1698	AF216283	AF216283 Acclimopol
33	231	12.6	1629	AF042332	AF042332 Oryza sat
34	228	12.5	11714	AE012928	AE012928 Chlorobiu
35	227	12.4	1383	ZMUT9669	ZMUT9669 Zea mays en
36	227	12.4	40065	SPBC16E9	SPBC16E9 S.pombe chr
37	226	12.3	840	AF216282	AF216282 Halorhodo
38	226	12.3	1080	AY113031	AY113031 Arabidops
39	226	12.3	1302	AF375397	AF375397 Arabidops
40	226	12.3	1497	AF045570	AF045570 Zea mays
41	226	12.3	97554	AC009978	AC009978 Genomic s
42	225	12.3	6085	A60304	A60304 Sequence 4
43	225	12.3	6085	AR144763	AR144763 Sequence
44	224	12.2	1094	D89131	D89131 Schistosach
45	223.5	12.2	1267	NT071108	NT071108 Nicotiana t

RESULT 1

ALIGNMENTS

JOURNAL
REFERENCE
AUTHORS

Unpublished
2 (bases 1 to 1047)

Kim, C.J., Chen, H., Cheuk, R., Meyers, M.C., Shinn, P., Banh, J.,
Bower, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D.,
Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Katlin-Neumann, G.,
Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,
Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T.,
Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C.,
Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinzaki, K., Davis, R.M.,
Theologis, A. and Ecker, J.R.

TITLE
JOURNAL

Direct Submission
Submitted (14-Mar-2002) Salk Institute Genomic Analysis Laboratory
(Signal), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the
sequencing and annotation of the RFL CDNA. (RFL CDNA: "RIKEN
Arabidopsis Full-length CDNA") : Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinzaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RFL CDNA. (RFL CDNA: "RIKEN
Arabidopsis Full-length CDNA") : Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinzaki, K.

Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinzaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES
source

Location/Qualifiers
1..1047
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/clone="U13225"
/note="This clone is in pUNI 51
ecotype: Columbia"
1..1047
/note="gamma-tocopherol methyltransferase"
/codon_start=1
/product="AT1G64970/F13011_27"
/protein_id="AA190941.1"
/db_xref="GI:19699150"
/translation="MKATLADESLTSLPYRTNSFGSKSLFRSPSSSSVSMTT
RGNVAATAATSTELARKGIAEFYNETSGIWEIWDHMHGFDSDSVOLDSGHK
EAOIRMIESTLRAGVTDEEEKIRKVVVDCGIGSSSRYSKSGAECIGITLSPV
OAKRANDLAASOLAHKASFOVADALDOPEDCKFLVSMESGEMHPDKAFVKELY
RYAAPGRITITVWCKRNLNSAGEALQDPDONLIDICTFYIPACSCDDYNNILDS
HSLODKADMSBNVAFWPAVYKRLTKWGLVSLRSGKSKTKGALIMPLEIGYRK
GVYKFGITRQKPL"

CDS

BASE COUNT 267 a 223 c 269 g 288 t
ORIGIN

Alignment Scores:

Pred. No.: 4..22e-92 Length: 1047
Score: 1154.50 Matches: 231
Percent Similarity: 75.43% Conservative: 30
Best Local Similarity: 66.76% Mismatches: 65
Query Match: 63.09% Indels: 21
DB: 8 Gaps: 4

US-09-857-613a-28 (1-350) x AY090280 (1-1047)

OY 8 ProthrlleScysIleHsIleThrPheargSerGlnSerProargTrhPheala 27
DB 53 CCAACCTCTTCTTGGGTCAAGTCAAGCGCTTCTTT-CGGTCCATCTCTCTCC 111
OY 28 ArgIleargVal---GlyProargSerTrpAlaProIleargAlaSerAlaIaSerSer 46
DB 53 CCAACCTCTTCTTGGGTCAAGTCAAGCGCTTCTTT-CGGTCCATCTCTCTCC 111

DB 112 TCAGTCTATGACGACAAACCGGTGAAACGTGCTGCGGCTCTCTACATCCACT 171
OY 47 GUAAGGlyGluIleValleuGluGlnTyrProLysAspAspLysLysLeuGln 66
DB 172 GAG-----GGGCTAAGA 183
OY 67 LysGlyIleAlaGluPheTyrAspGluSerSerGlyLeuTrpGluAsnIleTrpGlyAsp 86
DB 184 AAGGAATAGCGGATTCATCAATGAACTCCGGTTTGGGAAAGATTTGGGGAGAT 243
OY 87 HisMetHisIleGlyPheTyrAspSerAspSerThrValSerLeuSerAsp-----His 104
DB 244 CATATGCATTCATGCGCTTTATGACCCGTATCTTCTTCACTCTGATTCGTGAC 303
OY 105 ArgAlaGluGlnIleArgMetIleGlnGluSerLeuArgPheAlaSerValSer---Glu 123
DB 304 AAGGAACCTAGATCCGTATGATGAAAGCTCTCGCTTTCCCGGTATGATGATCA 363
OY 124 GluArgSerLysTrpProLysSerIleValAspValGlyCysGlyIleGlyGlySerSer 143
DB 364 GAGGAGGACAAAGATTAAGAAAGTATGATGATGATGATGATGATGATGATGATGAT 423
OY 144 ArgTyrLeuAlaLysLysPheGlyAlaThrSerValGlyIleThrLeuSerProValGln 163
DB 424 AGATATCTTCTCTTAATTTGAGCTGAAATGCAATGCAATGCAATGCAATGCAATGCA 483
OY 164 AlaGlnArgAlaAsnAlaLeuAlaAlaAlaGlnGlyLeuAlaAspLysValSerPheGln 183
DB 484 GCCAAGAGACCATGATCTCCGCGCTGCATCATCATCATCATCATCATCATCATCAT 543
OY 184 ValAlaAspAlaLeuGlnGlnProPheSerAspGlyGlnPheAspLeuValTrpSerMet 203
DB 544 GTTGGCGATCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 603
OY 204 GluSerGlyLysLysMetProAspLysAlaLysPheValGlyGluLeuAlaArgValAla 223
DB 604 GAGAGCTGTAGCATATGCTGACCAAGCCCAAGTTGTAAAGGTTGATGATGATGATG 663
OY 224 AlaProGlyAlaIleIleIleIleValThrTrpCysHisArgAspLeuGlyProAspGlu 243
DB 664 GCTCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 723
OY 244 GlnSerLeuHisProTrpGluGlnAspLeuLysLysIleCysAspAlaTyrTyrLeu 263
DB 724 GAACCTTTCGACCGCTGCGGACCAAAACATCTTGGACAAATCTTGAACCTTCTATCTC 783
OY 264 ProAlaTrpLysSerThrTrpAspTyrValLysLeuGlnIleSerLeuGlnAsp 283
DB 784 CCGGCTTGGTCCACCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 843
OY 284 IleLysSerGluAspTrpSerArgPheValAlaProPheTrpProAlaValIleArgSer 303
DB 844 ATTAAGTGTGCGATTTGCTCAGAGAACGATGATGATGATGATGATGATGATGATGAT 903
OY 304 AlaPheThrTrpLysGlyLeuSerSerLeuLeuSerSerGlyGlnLysThrIleLysGly 323
DB 904 GCATTATCATGAGAGGCGCTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 963
OY 324 AlaLeuAlaMetProLeuMetIleGluGlyTyrLysLysAspLeuIleLysPheAlaIle 343
DB 964 GCATTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1023
OY 344 IleThrCysArgLysPro 349
DB 1024 ATCACTTGCAGAGACCA 1041
RESULT 3
AY049258 1303 bp mRNA linear PLN 09-AUG-2001
LOCUS AY049258
DEFINITION Arabidopsis thaliana At1g64970/F13011_27 mRNA, complete cds.
ACCESSION AY049258
VERSION AY049258.1 GI:15146233
KEYWORDS F1L_CDNA.

QY	324	AlaLeuAlaMetProLeuMetIleGluGlyTrpLysAspLeuIleTrpGlyAlaIle	343
Db	1031	GCATTGACAAATGCCATTATGATCTTGAAGGTTACAGAAAGGTCATTTAACTTTGGTATC	1090
QY	344	IleTrpCysArgLysPro	349
Db	1091	ATCACTTCCAGAACCCA	1108
RESULT 4			
AX417728		1350 bp	DNA
LOCUS			linear
DEFINITION	Sequence 19 from Patent WO0231173.		PAT 18-JUN-2002
AX417728			
ACCESSION	AX417728.1	GI:21522874	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM	thale cress. Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.		
REFERENCE	1 Ethel, M., Kunze, I. and Geiger, M. Improved method for the biosynthesis of vitamin E Patent: WO 0231173-A 19 18-APR-2002; SUNGENE GMBH & CO KGAA (DE)		
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
Source	1..1350 location/qualifiers /organism="Arabidopsis thaliana" /db_xref="taxon:3702" 63..1109 /note="coding for gamma-tocopherol methyltransferase" /codon_start=1 /protein_id="CAD35425.1" /db_xref="GI:21522875"		
CDS	/translation="MKATLAAPSSLTSLPFRNSSFQSKSLFRSPSSSSVSMTT RGNYVAANAATSTALRKGLAEFYNETSGLMEETGDMHGFDPDSVOLSDGHRK EAQIRMEIESLRFAGVTDEEEKIKIKVYDGCIGGSSRYLAKFAECIGITLSPV OAKRNDLAAASLSSHKRSFOVALDOPEDFGKFDLWMSGEHMDKAKFYKELY RVAPRGRIITVTGCHRLTSAGEFALQWQONLIDCKITPTYLAMCSTDYVULOS HSLQDICKADMSENVAPFPWPVATRLATWKLVLSLRSGMSTKIGALTPLMIEGTYKK GVLFKGIITCKNPL"		
BASE COUNT	374 a 278 c 314 g 384 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.07e-91	Length:	1350
Score:	1151.50	Matches:	230
Percent Similarity:	75.43%	Conservative:	31
Best Local Similarity:	66.47%	Mismatches:	65
Query Match:	62.92%	Indels:	21
Db:	6	Gaps:	4
OS-09-857-613A-28 (1-350) x AX417728 (1-1350)			
QY	8	ProThrIleSerCysIleHisIleHisThrPheArgSerGlnSerProArgThrPheAla	27
Db	115	CCAACTCTTCTTGGCGCTCAAAAGTCATCGCTTCTTT--CGGCTTCGACCTCCTCCTCC	173
QY	28	ArgIleArgVal--GlyProArgSerTrpAlaProIleArgAlaSerAlaIleSer	46
Db	174	TCAGTCTCTATGACGACACACGCGTGAACAGTGGCGTGGCGGCTGCTCAATCCACT	233
QY	47	GluArgGlyGluIleValLeuGluGlnLysProLysLysAspLysLysLysLeuGln	66
Db	234	GAG-----	245
QY	67	LysGlyIleAlaGluPheTrpArgSerGlnSerGlyLeuTrpGluAsnIleTrpGlyAsp	86
Db	246	AAAGGATACCGGAGTTTACATGAACACTCGGTTTGTGGAAAGACATTTGGGAGAT	305
QY	87	HisMetHisIscLysPheTrpAspSerAspSerThrValSerLeuSerAsp-----His	104

[illegible]

Stanford University, 855 California Avenue, Palo Alto, CA 94304

4 (bases 1 to 118335)

REFERENCE
AUTHORS

Federspiel, A.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F., Altai, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P., Gonzalez, A., Kremenetska, I., Kim, C., Lenz, C., Li, J., Liu, S., Luros, S., Schwartz, J., Shinn, P., Tortum, M., Wyotskaia, V., Walker, M., Yu, G., Ecker, J., Theologos, A. and Davis, R.W.

Direct Submission
Submitted (11-JUN-1999) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304
USA

REFERENCE 5 (bases 1 to 118335)

Reuensispiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F., Altafi, H., Nguyen, M., Lam, B., Southwick, A., Ecker, J., Theologos and Davis, R.W.

Direct Submission
Submitted (28-MAY-2000) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304
USA

COMMENT On May 4, 1999 this sequence version replaced gi:4678192

FEATURES
SOURCE

Location/Qualifiers
1. .118335

```
gene      1092..2311
          /gene="F13011.1"
CDS       join(1092..1225,1474..2311)
```

gene	8107. .9734
/gene="F13011.3"	
join(8107. .8140,8222. .8358,8440. .8489,8562. .8893	
8959. .9062,9144. .9219,9304. .9734)	

gene

CDS

```

/genre="F13011.4"
Complement(join(10308..10729,10802..10883,11042..11255
11208..11384,11461..11617,11723..11956))
/genre="F13011.4"

```

```
/note="Location of ests E6H3T7 (gi|2414136), 62B10T7
/contig=1031264, 33081567, 6211317083 000 [location] +C
```

membrane related protein CP5 [Arabidopsis thaliana]"

```
/codon_start=1
/product="Putative membrane related protein"
```

```

/protein-d="AAD38248.1"
/db_xref="GI:5042409"
/translation="MNGLSIVPTFTHSSGQSDLIYFAEALMTVAAGVILVGVWVRRP
MAIDLSLINSKPEFLNLPLSITLTHSHVNVLSJSKRTGCVDDDEPHLKLVAWVRE
KWDGPDVLOMDDNPTPTSYQAWVRDENGPOYRSREIADPTPMVRFPEFDDDE
BSKDDMLLYSTLEKCRPTQYQAWVRDENGPOYRSREIADPTPMVRFPEFDDDE
GVQVYSVPQNDKPRDVLIVYSSMCIAVEKSKDGGSTCEVLLFHEDDGIIPMKIR
LGVRCMGCAVKKTEGRLAYQQAQAAGALSSPSIAHNTKYSAEFPMNRRSIAAE
VYDGPCTGKNIRLLIVYGCAIALACTLDKRLKNAVIFGVARRRANRGRK"
/size=5717

```

gene
CDS
13932..13971
/gene="F13011.5"
join(13535..13647,13808..14022,14535..14734,14807..15330,15402..15717)
/gene="F13011.5"
/note="Identical to Tubulin Alpha-1 Chain (Arabidopsis thaliana) (g11135391) Location of ests 139J20T7 (g112762968), 214C14T7 (g11159108)."
/codon_start=1
/product="alpha1 tubulin"
/protein_id="AAD38249.1"

/do,tiel-ig:30g2410
 /translatio="NREIISHIGAGIOVNSCMEILCLEHGIOPOCOTMSDSTVCA
 CHDARNTFSETSQCHVRAFLVDLEJPEVITDEVGTNYROLFHEBOLISEREDANN
 FARGHTVTEIVDTCLERKLADLCNGLOJFVFNVAVGGTSGGLSILLERAVN
 FCKSLGCTIYPSQVSTAVAEVPEYNSHISLLEHDVAVLLDNEIAYICRSIL
 IERPPYSNLNLISQITISLTSLRFGDILINDIVETFOINLPIYRHEFMSIAPVY
 SSAKAYHDEQSVPEITTSVFEEDNSMMAKCDPRHGYKMCCLMYRQDVVAKVYNVAIAA
 ISAKRIQJVDVNCPTGFCKGJINYOCPVYVPGDGLVORACVAMISNNNAVAEVSRI
 HFEDLMYSKRAEFVHWYEGMEGEGESSEARBDLALLEKDYEEVGEAGEADDEDEDEY

```
gene complement(16208..17133)
    /gene="F13O11.6"
CDS complement(join(16208..16338,17040..17133))
    /gene="F13O11.6"
    /note="Unknown protein; Location of ests 247E21T7
(gil2413341)"
    /codon_start=1
    /protein_id="AAD38250.1"
    /db_xref="GI:5042411"
    /translation="MAAEKAAATAAIVKKWDVPEDDDFEEFFETNEPWLAKPEVKRVS
```

gene
CDS
OMEDDDDDVNDVDSRSLRKLENTDKK"
complement(17641..19522)
/gene="F13011.7".18251,18334..18613,19281..19522
complement(join17641..18251,18334..18613,19281..19522)
/gene="F13011.7".18251,18334..18613,19281..19522
/note="313 identical to Glucan endo-1,3-beta-glucosidase
precursor (1-3)-beta-glucan endonhydrolase) [Trillium
aestivum] (g111706551)".
/codon_start=1

/product="Similar to glucan endo-1,3-beta-glucosidase
 precursor"
 /protein_id="AAD38251.1"
 /db_xref="GI:5042412"
 /translation="MATHTKLPKPTVQMLDNNINKVKLFDADEFTTMALAGSGLEVV
 VAPNOLAKMTSYDRAKWKNRVNRYRNFEDGSVNTFEAVQNEPFLSYNQSFLNLSQ
 FPLAIIQALNIEAGLGNSGVKATPLAAVYVDSGASNPVPSGSRPDIIGCMOIVDVS
 FLKNNAPLTINITYPLSLINDNDPELAFDGAEPIDINDIDTNTNPPDIDILVAS
 SKRAVHGMDPIYEVGMEGTEDGKHAANGASAFRNGCLPLRGINCKTPPTPTLYE
 YLGLGFLDEDKASTAPSPFERHMHIFEKDQOPKPIDLSGCGSKFLGKAAVPTPLNN

WCETPEAKDITJLKAANIDYACETSDCTALGYSSCNTLDANGNASYAFENNEFOVRNQ
DESACYFOGLATITTONISOGCCNPPIQIVASSASSPSCSSYSLVVLIVNELLSGMMF

```

/productSimilar to glucan endo-1,3-beta-glucosidase
precursor"
/translation_id="AA038251.1"
/db_xref="GI:5042412"
/translation="MAHHKLEPKTVQGLKDNKNTKKKLPADETFMGALSGSELYN
VALPNOLKVTSTSDKRLQVRRKRVTRNDEGVNITFVAAGNEFKSPISYGS
PALAIQNALNDEGLNSKYATPLADVADYNSVNPAGRRPDIICOMQIYDV
FLKNNAPITNTPELSTGDEQEPPLNTAFEDCAEPLINDGIDTNEADNPTLYE
SLKAVGHGMPIITVTEGMEQGHKNAAGSAVRPPNLTPLTGNCTPLRPYIEV
LYGLLEDEDKSTIAPGPERHNMGIETKEDGQKFPITLDSGQSQKFLJQAQNVPLPNRK
WCFENPACAKLTIKAAINDIACFTSDTALYAGSSCNTDLNAGNASTAFENFQVKN
DEACACFOGIALTTITNIDSOCCNFPIDYIIVASSSCFSCSSYLWLVILNGLCMHNE

```


source
1. .343550
/organism="Nostoc sp. PCC 7120"
/db_xref="taxon:103690"
/note="synonym:Anabaena sp. PCC7120"
gene
complement(94. .411)
/gene="a111719"
complement(94. .411)
/note="ORF_ID:a111719
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB73418.1"
/db_xref="GI:17130809"
/translation="MRKKSATLEAVHETAEDLHKAGLMQTLREFEHLCPPIEPL
EPQAIKEIRRESSQVAVFARIINISTVQKWEIGOKRPSGASLKLHLVKNNGLS
VLY"
gene
complement(504. .782)
/gene="a111720"
complement(504. .782)
/gene="a111720"
/note="ORF_ID:a111720
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB73419.1"
/db_xref="GI:17130810"
/translation="MKIKNRTFDMARKEGKLNLSICNAVNEMAPGLYDADGGGLF
KRIAKPGKSGCFRTLATNNEDEVGFLEFLAQKTAVALTHMKKTL"
935. .1507
/gene="a1r1721"
935. .1507
/gene="a1r1721"
/note="ORF_ID:a1r1721"
/codon_start=1
/transl_table=11
/product="transcriptional regulator"
/protein_id="BAB73420.1"
/db_xref="GI:17130811"
/translation="MRHKDNKNOAICDAAIELITANGPADTSMKTAKANVSPATI
YVFEENKEKLKILYLFVKQISAEILGVNONLSVEKAFKMNNYKKAIKPVPF
APTEOFANSPMDVRICKDESLYEKPLMDLFPNGKEKVKKDIPLEIFTAFTAPLGLG
LKEHFSGAILDEKMLETTYKIAMDAVTN"
1569. .2504
/gene="a1r1722"
1569. .2504
/note="ORF_ID:a1r1722"
probable oxidoreductase"
/codon_start=1
/transl_table=11
/protein_id="BAB73421.1"
/db_xref="GI:17130812"
/translation="MKNEKNAENIPQGRKRLAVYSSSGICYETARVLANKOASVI
IAVNLDKGKALAKILQONKADADVKLMELDLNLSVKNFAENFRKNYLRLLDLIN
AGVMIPIYSKTTQFELQFGTNLGHFALGOLLEFISTEGSRIVVSSGANMGI
DFEDLMDQESYAKRAVGDCKLANLYETELDRKDKONGIDILVTASHGWTATELO
RTAGGIVKYLVADODITMGALPTLAAIEAGIKGAELVGPNGCFMGRGYPKIVESN
ELSKDQALAKKLVVSEKLLTDVKEFEFNKKAOSAGK"
complement(2701. .3462)
/gene="a111723"
complement(2701. .3462)
/gene="a111723"
/note="ORF_ID:a111723
probable aldolase"
/codon_start=1
/transl_table=11
/protein_id="BAB73422.1"
/db_xref="GI:17130813"
/translation="MSVFSRPLPVPFNSIEERLHRKORLAAAFRLGKFGFSGSG
HITARDPEFDHFVMDPKYFGHIRYSIDLVLNRDGDVYKGOALLINQAPAIHSOIH
EARDVYAAAHANSIYKANSISGLRLDPTDQACAPYEDHALIDDTGVVLETSEBK
RLAETLGNKALILRNHGIITVGHIVDEAFAWYIISERSCOQLLAFAAGPKIIGHE

gene
complement(3764. .5167)
/gene="a111724"
complement(3764. .5167)
/gene="a111724"
/note="ORF_ID:a111724
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB73423.1"
/db_xref="GI:17130814"
/translation="MSSTSGVPGCLDULHPTTELIGLSOLITASLLVFLVIALGIA
IALMSFALRHNSQQTIFIGEMAIRAQILIRGLQHLALVILLVGGEFFLSSTLSNRH
HMDQAKIAQVQVSGDKLEQIAPOVRYVQDEPYTNTQVNGKIVKYNDOQVSRLLS
LACSQIOVKLEOSTDVQGRSAVRYADYADYKVYNRLKIDINFEFEIOPNGYSLNS
KYVERORTLEPVPNGNVSFAFLEGOETSFVTOAQGPRVYNSAGOLLSNFRLL
TTIANPQADAFAGIVPKERKADGRSTQVYIMDDVNSVKNPGRVFTNTPIDTQGITI
PRILLAPALFLMWILLIYLSLPSISLTNVAIASIYFACILITITTSYSRDQALNTM
ISLVLLFTWGLGSTRSASLAIICTAGAVVPLDLVLVPSQLTSLVAGLSTVWLA
VRHMYMNIWQPEDSKS"
complement(5492. .5902)
/gene="a111725"
complement(5492. .5902)
/gene="a111725"
/note="ORF_ID:a111725
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB73424.1"
/db_xref="GI:17130815"
/translation="MFLKPPSGLMVITASLTUNGIAITYSGRSIAEAKSNPEYLVQGA
DQGFPIALDLASINGTFTLVQOHGEGIAKRRLHAACSGRLFTKFSLYAANGKLT
SEDKTEQEIFPKPKPTDPAASMEIVCRVANHQKK"
6321. .6842
/gene="a1r1726"
6321. .6842
/gene="a1r1726"
/note="ORF_ID:a1r1726"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="BAB73425.1"
/db_xref="GI:17130816"
/translation="MKSQHLQATTLAGEVSQOYFKETALNAIAELQDFIDMCPDARE
VRKALAVKLVYQGYLDYDELOKILDVSRGSITGKQAYEEGIDGLRYSGRKSYLNA
KQOEVLISMLQTKDCWELGELVEYQLAFEYDVVESQSYDLPSEAGISMKKTKVNP
KADENAVAEKKRD"
6919. .7431
/gene="a1r1727"
6919. .7431
/gene="a1r1727"
/note="ORF_ID:a1r1727"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="BAB73426.1"
/db_xref="GI:17130817"
/translation="MWGDISGYWGSQDOELAVPVNERNKORYYGAVDYLGEIVLK
AYDAGSKNTINTLEVLANSPPQRLIIFPDGASYRSKRTIRFLDSVNSLPTEDQK
HCVFRAPNCPVONPIEDIWLOAKTWRRCALIPFSHLKWIIFEWIRHTTDFDPL
OMYGTYSKIR"
8064. .9401
/gene="a1r1728"
8064. .9401
/gene="a1r1728"
/note="ORF_ID:a1r1728
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB73427.1"
/db_xref="GI:17130818"
/translation="MTGNARNAHAFRLVSGNGCAFSGSRISLTSKEVYIGRDPSCQ
VLIDAMYRNVSRRAHVVRVAVSSVSKFSWVLCIDLNSANGTYLNGOQLRYGCGEILHAG


```

/gene="SB234M12.3"
/note="highly similar to Oryza sativa gene for putative
receptor kinase GB:AF128457 GI:4680335; similar to Oryza
sativa receptor kinase-like protein gene GB:AF119222
GI:4680488; similar to Arabidopsis thaliana receptor-like
protein kinase emb:AL138654 GI:6899879"
/evidence=not_experimental
complement(join(<12094..12725,12823..>14263))
/gene="SB234M12.3"
/product="putative receptor kinase"
complement(join(12094..12725,12823..14263))
/gene="SB234M12.3"
/codon_start=1
/evidence=not_experimental
/product="putative receptor kinase"
/protein_id="AA04320.1"
/db_xref="GI:22208505"
/translation="MTTLARLPPELLLAATAAANDNGNTTEAAMAALVNL
KSTPTDPTGLEAMSATSPPPCDASPPGVCYKSLVGRVLTNNLSGTFDGL
AKLRLHSVNLKHALSGPLASLGLRLALYLSNNSGPIPADVFAMRWLKL
YLDNNRITGLPADAIKAPRLMELIDRNOIDGPVFKLPESLKRNVSHNRSGTI
POSVAERYDASSFAGNPGICSGSDAACAAPALPPAPPTPAADYAMEETSV
FVYIGTILVILVSGAVMLRDERNSVATANDYAGSTTTGAGAGAGSGSKVA
DAASTTAAPRAGAVADMASSGSSSHGGSGSSGRNGEYVLDNEDITAFGLPDL
MKASAEVYNGTIGSAYKAAMRNGVTAVKRLRDMNRFRDEQHVHMLQLRHPNV
LPVGYHARKKEKLIVSEYMRGSLYILHGDQSNRVTIDMOGRVIRAVGVGLAF
LHERLGIAPRGLVSMGDGADFDAPPPEPHGNLKSNTILGADMEPLVDYGFPLVA
AQAOAMPAPRSPGCTRGVARSADYVCLGVILEVTGFRPSQYILSAGGDDVH
WAATVAVEGGERELVDPAITAAAGDAVRLRVRGHCASPEPRCPVAEAAWVEET
GSAAGAS"
12558..12578
/repeat_region
/rpt_unit=c99
complement(13145..13307)
/rpt_unit=c99
13490..13517
/repeat_region
/note="GC-rich"
14190..14216
/rpt_unit=c99
complement(<15173..>20442)
/gene="SB234M12.4"
/note="Similar to Arabidopsis thaliana gene for putative
sugar transporter protein GB:AY094465 GI:20453188"
/evidence=not_experimental
complement(join(<15173..15605,16042..16226,16382..17055,
17854..18423,18531..18650,19447..19501,19596..19649,
20293..>20442))
/gene="SB234M12.4"
/product="putative sugar transporter"
complement(join(15173..15605,16042..16226,16382..17055,
17854..18423,18531..18650,19447..19501,19596..19649,
20293..20442))
/gene="SB234M12.4"
/note="AT4g35300/F23E12_140"
/codon_start=1
/evidence=not_experimental
/product="putative sugar transporter"
/protein_id="AA094321.1"
/db_xref="GI:22208506"
/translation="MSMICKLNVAQEDLMGATGAGQDVQWNGKLEEVAVLCAH
RIMAEVSTHOEVCKMYFKFAGAAIASRLIDCEAYLRTEFGFIVCKRAVLDEM
STVFSAVVYISGYAMGDFALLLEAHNMKFEELLNPSLEGITLAASTGATVIT
IFSGALIMDLGRRALIVYSILVFGSGVIMLSPNIYILRLRILVSGSGLVFICP
ITITSETSPNNRGLCTMPQMFICTITSTCLITMLTMSPPNRIMICAFAPSVI
YFALLVYILPESPRWLADSKISEARVSLQRLRGKHDSGETAVIVDQVDTISBSAV
GTARAOSEFGTSATRWPRSTIYWDLSLVLGSHHNMSEGGSRNSFPFVSSE
SEPEHHNNEHNDGSDQOTREYSAGVNNGSLFASLISAAVEVANDANTVFSE
GSSYLRHGTGLAOEFMAISHDHDIEESEEHGLFALHOSAPRVESRGRQVYRL
SETADMKERVRILQDCTIRALCYGMLTQALQOSNGISCLLRAPILLQVQYSLSD
IGLSRHSITLITDQALILPCTRAMLPMDCGRVYLATPPLITLISYMSMSI
LENNMGFRAIVFHFALVFCSVYGLDIPILICSEIPTARARAFSCSLSWSE
GGLISAYCLPVMVSTIGLGACGIVAVLCCAPLFLFYRIPETKMLLELLEFLKS
ROEYVO"
complement(22125..22161)

```

```

/note="AT-rich"
complement(22424..22446)
/repeat_region
/note="AT-rich"
23019..23089
/repeat_region
/note="AT-rich"
23130..23172
/repeat_region
/rpt_unit=ta
23495..23526
/repeat_region
/note="AT-rich"
26812..26900
/repeat_region
/note="AT-rich"
37014..37038
/repeat_region
/note="GC-rich"
complement(37399..37462)
/rpt_unit=a
<37578..>44288
/gene="SB234M12.6"
/note="Similar to Nicotiana glauca plumbaginifolia ubp1 gene
emb:AJ272011 GI:6996559"
/evidence=not_experimental
join(<37578..37721,38745..38806,38998..39100,40241..40317,
40518..40579,40690..40777,40913..40988,41080..41110,
41205..41318,42137..42222,42364..42537,43066..43258,
43335..43537,43864..43887,43980..>44288)
/gene="SB234M12.6"
/product="putative oligouridyate binding protein"
join(37578..37721,38745..38806,38998..39100,40241..40317,
40518..40579,40690..40777,40913..40988,41080..41110,
41205..41318,42137..42222,42364..42537,43066..43258,
43335..43537,43864..43887,43980..44288)
/gene="SB234M12.6"
/codon_start=1
/evidence=not_experimental
/product="putative oligouridyate binding protein"
/protein_id="AA094322.1"
/db_xref="GI:22208507"
/translation="MHPRMKRAAQAQQLMQOALLQOQQAQPPRFCHHHPKGL
AAPQMEPIVSGMLPPGFDSTCRSVYGNISLQVDTYLADEVQSIGVEGKLRKE
KSSFGFIDYDRYALALAILSLNGRPYLGQPIKVMWATYSTREDTSGHNFVSDLC
PEVTDATLFAFSGYSTCSDAVMDQKGRGRGFGVFRMOQAOASAINDLNMKL
GNROIKNMATKGANAGAEKQIYDQKVDLNTSSESGKPNEDGESNPFYTVYG
NIPHEATNNDVHLFRHSGACSGSTIEVRTRDQFGFVRSTHEBAQAIDOMANGOLIG
GQOIKTTTNGYKRLAVESDPCEYLKWSRLVTLVYVASVFLLEGNDAVRO
GIDLLYASTAGFAVGCGLVGLVYLYLVAOSDPBTLRMKGAQACFACVYR
SKLCGSGVKECAKEGAIVAVARCTEMLGMSGSKPTGTSASPPLPAPALFSTVG
SPTDPLAVORMLNMTNPSIMGHALNQVALMGGRASQSIYDGFQGISMAAT
SNOQOQOQOQOQIMY"
37580..37668
/rpt_unit=c99
complement(39121..39142)
/note="AT-rich"
44247..44274
/rpt_unit=c99
complement(45252..45289)
/rpt_unit=c99
<45290..>46757
/gene="SB234M12.7"

```

```

Alignment Scores:
Pred. No.: 4.17e-39 Length: 108553
Score: 579.00 Matches: 160
Percent Similarity: 38.908 Conservative: 31
Best Local Similarity: 32.598 Mismatches: 61
Query Match: 31.648 Indels: 242
DB: 8 Gaps: 5

```

US-09-857-613a-28 (1-350) x AF527809 (1-108553)

OY 94 AspSerAspSerThrValSerLeuSerAspHisArgAlaInlIleArgMetIleGln 113

DB 94288 AATTCACCTGATTATATAATTGGTTGAC-----ATCGTTCGCAACTTTGA 94241

OY 114 GluSerLeuArgPheAlaSerValSerGluGluArgSerLysTrpProLysSerIleVal 133


```

gene
CDS
    complement(1482..2213)
    /gene="ho2"
    complement(1482..2213)
    /gene="ho2"
    /note="ORF_ID:tl11721"
    /codon_start=1
    /transl_table=11
    /product="heme oxygenase 2"
    /protein_id="BAC09273.1"
    /db_xref="GI:22295447"
    /translation="MTVLSLALREGTAEAHTLAHTAPMKCFVGLITSRLRLOFLAN
    LYFYVDTLEAASNGDAVAIYFPELNSGOLAODLAYEGEDMDQYTPLATATY
    YLSRLHDVAOSEPLLIASHYTRMGDLSSGOALSKTVHSQMLPEKGTAFYEPAL
    PTEBDKAFKORTATIDSGLDEEATARIYQENAFALNCNLMALESBLKTTIGD
    QVMSVNIENPSRHOHPREALAI"
    2486..3550
    /gene="tl11722"
    2486..3550
    /gene="tl11722"
    /note="ORF_ID:tl11722"
    hypothetical protein
    /codon_start=1
    /transl_table=11
    /protein_id="BAC09274.1"
    /db_xref="GI:22295448"
    /translation="MVAIAPRPHSGSVKTIITREMLTPREYTTDFEQAKLDSROE
    QLEMLKMRADNRHNFVRDSEKGTWDLIDQTRAFIDYLERSCVSESGFLRK
    OLSRLKLRNPDLAEIRFLMARDEARHAGLNMKMDFGHAMDGYLSKRTYTFPL
    PMWLYVYLSEKIGYMYRIIYRLHEKHPSEFYPLNYFERMCCODENRQDILKALI
    HSOPOLIOGSAKAMREFFLTIVATHTTLVHERAKFYEALGDATAFDREVIKATHE
    TAARTFSVLANVDHPEFYSLQRCEVELSNLQALAEATHPQKWLKALKPLHQALIAH
    LIRYLLPYNAGQENCTYH"
    complement(3347..3858)
    /gene="tl11723"
    complement(3347..3858)
    /gene="tl11723"
    /note="ORF_ID:tl11723"
    hypothetical protein
    /codon_start=1
    /transl_table=11
    /protein_id="BAC09275.1"
    /db_xref="GI:22295449"
    /translation="MLSRTPLIERLAAIGSDIYMDYAKKMLYLGDAKLAPLAEFL
    PLEQGMSLAQYTSVLQEVSVPLGGDQFLSLERLLIPKANQELIVELNRFQEELE
    S"
    complement(3963..4382)
    /gene="psad"
    complement(3963..4382)
    /gene="psad"
    /note="ORF_ID:tl11724"
    /codon_start=1
    /transl_table=11
    /product="photosystem I subunit II"
    /protein_id="BAC09276.1"
    /db_xref="GI:22295450"
    /translation="MTTLTGCPPIYGSGTGLSADTBEKYAITTSPKQVEMRL
    AGAIVMRGEGELVFARKEQCLAAQOLRPRIKINDKIRIFPDGETVLILHPQGVF
    PEKVNKREAVNSVRSIGONPNSQLKTKGKKPYD"
    4763..5398
    /gene="tl11725"
    4763..5398
    /gene="tl11725"
    /note="ORF_ID:tl11725"
    hypothetical protein
    /codon_start=1
    /transl_table=11
    /protein_id="BAC09277.1"
    /db_xref="GI:22295451"
    /translation="MYLLYLHGFASSSRSSKAQYFRDRFEELGQPLLPIDINQDFPH
    LTLRQIEQVEALLTPPEPITLIGSSFGGLTAAWLAEKHQVVOFLFLAFAFEPAIHW

```

```

gene
CDS
    complement(5383..6369)
    /gene="tl11726"
    complement(5383..6369)
    /gene="tl11726"
    /note="ORF_ID:tl11726"
    /codon_start=1
    /transl_table=11
    /product="delta(24)-sterol C-methyltransferase"
    /protein_id="BAC09278.1"
    /db_xref="GI:22295452"
    /translation="MSHLGLILVTVAVALVGLALTYLPPRYESARSVAESYDMT
    KDIILEFYMGCEHILHCHYGILPPPKDFQAKVDVFHEBMVWAGIDRLPPTGYLDVC
    GIGSSRIILARDYGTITISPEOVRARRELPALANRFOLDALALSPDASFD
    VVNSIENGPRMPKQCFARKELLVYKRGITLVYADNQRDDROQPLNFMRLITRQLL
    DONAHFRASIEGRAEALANTGLVAGVMTADTQETLPESMLISMOGIVRPEGLIRF
    GLPGLVSLREVPFELMLRIAFGCLCFQMFRAVRAEIPAVSLPEAPVNC"
    6619..8262
    /gene="rlda"
    6619..8262
    /gene="rlda"
    /note="ORF_ID:tl11727"
    /codon_start=1
    /transl_table=11
    /product="riboflavin biosynthesis protein Rba includes
    GTP cyclohydrolase II and 3,4-dihydroxy-2-butanone
    4-phosphate (DHBP) synthase"
    /protein_id="BAC09279.1"
    /db_xref="GI:22295453"
    /translation="MPVFDISIALRGEYIVVDENRENEGDILGAERVTPA
    MINEVAHAGLCLAMEGRDELPLMTVTTSDNSQTAFTYSVDGARWYTTGI
    SAEBRARTIALIDPSTPODLRRPGHVPFLRSPPGVLRKAGTEAVALUTRLAGLY
    PACYICETIONPDGMSRLPOLMEYARHOLKITISIALISRLQHEFRVREAVAKLP
    TAGEFPIYICRNSLDSSEHAIYKGPATFGQPVYRVHSELTDAIGSLACDR
    MOLOALKMTNAGRGVYVLRQGRIGLVNKRATSLDDGLDITYEAMHLSFEPD
    LRNVGVAQILNDGVKQIRLTNNPRKLAGLGYGLEVDVDFPLLEAVPYNPPYL
    TKAEKILGLIOTYLTIVAFRMQSOGLAADRERLEKRLHLAGVHLVREEARPYA
    QAIKPHDPLIVHGFQDPHYADRWQYQSPHPYLAVOTLLRQCCMPTLKGFELVA
    TGTDPMLNLPMHDLDRQAYTQOSPSEMRPHLIYSNAAG"
    complement(8267..8455)
    /gene="ts11728"
    complement(8267..8455)
    /gene="ts11728"
    /note="ORF_ID:ts11728"
    unknown protein
    /codon_start=1
    /transl_table=11
    /protein_id="BAC09280.1"
    /db_xref="GI:22295454"
    /translation="MDVTRDKQCHNSVSIYIWKSLAPLNKRVTRHPEKGVNLGVVP
    TPOSOTCCOYSOTSIGAF"
    complement(8456..9043)
    /gene="tl11729"
    complement(8456..9043)
    /gene="tl11729"
    /note="ORF_ID:tl11729"
    similar to heli protein
    /codon_start=1
    /transl_table=11
    /protein_id="BAC09281.1"
    /db_xref="GI:22295455"
    /translation="MMDCPLPIVIVPQMRAPHPNLTDSSALHLMLTLPPDVRGYI
    LRATLRVOPNLGERPLPRAGGKPYVINGLEFNWSHGSLAVLASGAAVGVNVELL
    RLCPQRAISRFRFGAALQDRILEGGSRSTLOANTTYEAMLKAGDIGVWORTAOSLE
    HWVASFPVGDAIASVVLTTPPOCFEFPREAVA"
    complement(9050..9559)
    /gene="tl11730"
    complement(9050..9559)
    /gene="tl11730"
    /note="ORF_ID:tl11730"
    hypothetical protein
    /codon_start=1

```


/db.xref="GI:17131112"
/translation="MKTPTLDMNEAVEGCAALAGITLGYGRVPLPOLRRCGSRD
GSKASNMVAARTYGLLEAKFKELKQVQQLPPYIVENNHNFEVVEGFGHQVYIN
DPATPRHVSIFDEEGYGLVLMPEGAEFGQGRKPMIASLMSRLGAGSALVYC
IVAGFELTIGLIVFVSQIEVDLILVGRMLLPHLLAMAIAAIILOSGLTLRLRY
LRLKILRLAVGSSREHMLILRLPYSEVQRAGELSNRTTINDOVLAQGLATTI
IDTLMLVFLAVMVOYDWYLTLMVYFAFVAVNLTLQMTISRRIDANORLIOEYGAAG
ISIALLOSIELTKASGLSEDFSSRMSGYTKRLNSOELGITNOFELSVLPTLSA
MLLIVGGFRVMDHSLIGMLIAFGGLMSGFOQPVNNLVNFGTTLQLEGNLIRLDVA
LDNPGVEGKTSSSSLQALPKLQGVYELKNTTFVSRLEAPLLENFNLISIKPGQVVA
LVGSGSGSKSTIAKLVSGLYCPMTGEIYFDSQPREISLQLTNSVAMVEDLILFEG
TVRDLNLTMDTVPDPKNLMRACODAAIADVIFMSGGYDAELIEGANLSSGDRLE
IARLVNNSILIMDEATSLADAFETEKIIDQNLRRRCGCIIVAHRLSTIRCCDIIV
LERKVVQAGTQOELMQVEGATSRILRTG"
/gene="all12022"
/complement(6028..7395)
/gene="all12022"
/note="ORF ID:all12022
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB73721.1"
/db.xref="GI:17131113"
/translation="MOVSPKSMPLVNLGSLVGVAFIMSTIGSTPIVEGSGVLIY
SNVPLQSKSAGQIMANVKNVDYIKKQVYLTQAEIRKQLQOORGLTOLESODK
AVGSLQGRLEDEERSQOQROYLRQLLELAIPLTKNTSVIQOQROGLQRLQ
QAQALTAVFLRRMKIROELFOKEGASIGSEALKAQOEYLNQLEKADIAQOAKELDV
ETQOEDKRENTLVLADLOSQLOQSOASVAOODLENTARNKEIOEVRKEISLE
LQLSNSOITISQHSRILEITLPGOVNAGTRLATIEANPQSKLGVGYTPVAVG
KIQPGMSIOTIPQYKRRERFGIVGNVTSISRFPTKEAANEVNSVELTIVSQQ
GLIYFVSHLELDANTPSGCKMSSSTGPHLNISSGTTTYVRVVEERAPITFVPLTRS
ASGTY"
/gene="all12023"
/complement(7472..9502)
/gene="all12023"
/note="ORF ID:all12023
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB73722.1"
/db.xref="GI:17131114"
/translation="MSDYCLVMPYAVERBSIALGLKARIQAKLEPTIYPIQF
AEICGLDYAMISENFIETFLSEMTFSGVLEPERPNHSEYFDLISPCFKSTENLPD
RNLEKAFIAVREQAATFIESVAQSVANLQPRILIGCSSTFOHCASIALKRYRELAPE
ITTFMGANGECAMGLATHTESVWDFTISGEGDLFNEFCYQLDGRVEDPSILPY
GVISHSRSTKTSIAAPRASVQNLDOVPIPDYDIPOTIAHSTLSPYIOEGLIETS
RGCMWGGKHCFGCLNGSGIYBRSKSDRYAKERTQOCERGLKRFQVNDITLDMN
MNTVLPVLAQDEANITYETKANLRWEHLQQLAAGVRCIOPGIESMHNSTLKLANK
GNTAMNVOQLKWARELGIIVFNILADPGSDSDMYEMEMFLPLIAHLQPPSGASE
IRYDRFSPIYERPSDMGNLSELRPKYVYPLAPILSELAFEPDQNNLVRYATSE
NEOTSEKTERPLENLKLMVQNMRLQSLYKHKPEYLPILSMNDSGLTITIDSR
PCATESLFLBGLAVWYKVCDRALTYREVASEIKKPEYLPVPMDEIRPIYSELARER
ILLDIDGRLLSALGSCSPYISVSELPAGYIDLHEFRHNSKSEKLLPR"
/gene="all12024"
/complement(9616..9909)
/gene="all12024"
/note="ORF ID:as12024
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB73723.1"
/db.xref="GI:17131115"
/translation="MSLVTWKLEFEKIIIVAGODESFRKQALVNPSTLEKGINLP
SSIDVNLVETTRGNLSLOALPNSBODTISEALESEVAGRISGEVSYIVKVF"
/gene="all12025"
/complement(10195..13482)
/gene="all12025"
/note="ORF ID:all12025
similar to Mrsd protein"

/codon_start=1
/transl_table=11
/protein_id="BAB73724.1"
/db.xref="GI:17131116"
/translation="MNVSKOSILIKTIERSNSISERLSSSEFIENDROINANONKLEK
WYOFIAKNOHNEFEKRLMDGLDTPTTICVGLNVLADBEKILPNVETLKGMOVATE
VIPENSNKFLQOHRHLPDRERFPEELIVPVDAKKLLIESAGSSTYLLSQRNHIQ
MERSLLEHVTGICISPLEDEFSIFRSLSSEPLVRFINOLQSGSKQOYTDVKNLKKO
GLISFQESVYARLILSVDFVAVTEFELSRLASDSQKLOLTFQPEQKIGOVVYQ
LAMSDFHNRGAVIIVIKPASGKLVYKKRSLCLEEAYDFLDMINOKINLPILKLI
ILNSVGNHFEAFALPCODQKVKHYORAGVNLIVLTKGNDCHENLACGQOPV
LVDTLELHHRFWLSKDDADKASINELQDSVGTGILPGMOIIPYCEOTELIKIDFS
GLGGEDEMPFRANKKHINTDSVAVIYRETTKLLPKNRPGENITDSLNHNSKELI
DGFROMYFLOQKKEELASDSPTIAFSNOKRVLRINITYVASILQNSLNHNSKELI
VEHSIQDLDSRAFLSSSDKRLPMLPLAEKQALBQDLPYTFAYSQNSVNDIISSEOT
INKEFLSSSYDDVIAHLROMDADLAKOISLRGSLYGLCTEQLNSLPLEPNLKM
IYAPIERGMLOQAIAIOOLEQOARCKQNSVWIGWGYFEAORPOLOPVGNSLYD
GCCGSLFLAALASVYNNAREGELSLASLOSKRTLODPHIDLOOKLIGOMIGGATG
LASIYVALVRISSEFGEVBLIODNOQSLASMTINATYNOAGVNTGYAGTITSLAL
YKATREPATLQQAISMGNTLDTRTVTVDEVOAMNABEKVLGFSGAAGIATLQ
LYAATDEPIFASAREAITLBOERLFSNSCODLDVADISNLSSTTECSGLIGIS
VARLCTNLIDPDEIROELINLHTIHSSIQVLDNICCGNFSHTEMLIVASQKLSHP
ELLETPANNRATQVLSRANOIGYFOLLPHLPADVVSFGCGQIAGIGYELLRLSCDDL
PSILIMO"
/complement(13877..14227)
/gene="all12026"
/complement(13877..14227)
/gene="all12026"
/note="ORF ID:all12026
similar to nitrite hydratase"
/codon_start=1
/transl_table=11
/protein_id="BAB73725.1"
/db.xref="GI:17131117"
/translation="MSBOTKPKREIAOIIYOAWKDETYROELLNNSKAVIEREPAIO
LPBEINHVAVEENDSNFVFIIPARNLEDELSEBQLEAVAGSLGDLFTFEVETTYDR
LKDRCKEIKAYVD"
/complement(14385..14720)
/gene="all12027"
/complement(14385..14720)
/gene="all12027"
/note="ORF ID:all12027
similar to nitrite hydratase"
/codon_start=1
/transl_table=11
/protein_id="BAB73726.1"
/db.xref="GI:17131118"
/translation="MSBOTKPKKVDENAOIIYOAWKDEYROELLNNSKAVIEREPAIO
LPBEINHVAVEENDSNFVFIIPARNLEDELSEBQLEAVAGSLGDLFTFEVETTYDR
LKDRCKEIKAYVD"
/complement(14854..15042)
/gene="all12028"
/complement(14854..15042)
/gene="all12028"
/note="ORF ID:as12028
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB73727.1"
/db.xref="GI:17131119"

Alignment Scores:
Pred. No.: 3,856-22 Length: 334520
Score: 392.50 Matches: 102
Percent Similarity: 51.70% Conservative: 50
Best Local Similarity: 34.69% Mismatches: 101
Query Match: 21,458 Indels: 41
DB: 1 Gaps: 12

US-09-857-613A-28 (1-350) x AP003588 (1-334520)
QY 69 IleaIagluPhetYrAspGlu-----SerserGlyLeuTrpGluAsnIleTrpGlyasp 86
DB 127499 GTAGCAATTCTCAGACGACGACTGACGACGATATTATTAGAGTTTACTGGGCGCA 127440

QY 278 SerLeuSerLeuGlnAsp-----IleLysSerGluAspTrpSerArgPheValAlaPro 295
 DB 670 GCCACGGGTTGGTGGAGGCCAGGAGTACTGCTGCTGATGAGATGACCGACCTCC 729
 QY 296 PheTrpProAlaValIleArgSerAlaPheThrTrpLysGlyLeu 310
 DB 730 GCTTGGTTGGATACCATT-----TGGCAGGGCATT 759

RESULT 15
 LOCUS AX073657 957 bp DNA linear PAT 06-FEB-2001
 DEFINITION Sequence 1 from Patent WO0104330.
 ACCESSION AX073657
 VERSION AX073657.1 GI:12710078
 KEYWORDS
 SOURCE Synecocystis sp. PCC 6803.
 ORGANISM Synecocystis sp. PCC 6803
 Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
 REFERENCE 1 (bases 1 to 957)
 AUTHORS Herbers, K., Badur, R., Kunze, I. and Geiger, M.
 TITLE Identification and overexpression of a dna sequence coding for 2-methyl-6-phytylhydroquinone-methyltransferase in plants
 JOURNAL Patent: WO 0104330-A 1 18-JAN-2001;
 Sungene GmbH & Co. KGAA (DE)

FEATURES
 Source Location/Qualifiers
 1..957
 /organism="Synecocystis sp. PCC 6803"
 /db_xref="taxon:1148"
 1..957
 /note="unnamed protein product"
 /codon_start=1
 /transl_table=1
 /protein_id="CAC28415.1"
 /db_xref="GI:12710079"
 /translation="MPEYLILPAGLISLSLAIAAGLYLTARGYQSSDVANAYDQMT
 EGGILEYMGDHIHGHYGDHPVAKDFOSKIDFVHAMAQMGDLTPGTTVDVGC
 GIGSSRLIAKQVGNFTGISPOQVRAETLPPDYTAKEAVDAMALSPDSSFD
 VYMSVEKGFPHRDKAVFAKELLRYRKPGCILVADMRDRDPLNFWKRYVAKQLL
 DQMSHPAASLEGFENLEATGLVEGOVTDVTPVLPAMDITWQGITRPGWLOT
 GIRGFISVREVPAILLEMLRIALFVGLCRFGMEKAVRNATQA"

BASE COUNT 203 a 244 c 268 g 242 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2,46e-25 Length: 957
 Score: 392.00 Matches: 92
 Percent Similarity: 51.76% Conservatve: 40
 Best Local Similarity: 36.08% Mismatches: 87
 Query Match: 21.42% Indels: 36
 DB: 6 Gaps: 10

US-09-857-613A-28 (1-350) x AX073657 (1-957)

QY 69 IleAlaGluPheTyrAspGlu-----SerSerGlyLeuTrpGlnAsnIleTrpGlyAsp 86
 DB 106 GTGGCCAGCGCTTACGACCAATGACAGAGAGCGCATTTTGGCAATATTACTGGGGCGAC 165
 QY 87 HisMetHisHisGlyPheTyrAspSerAspSerThrValSerLeuSerAspHisArgAla 106
 DB 166 CATATCCACCTCGCCCATTTAT---GGCGATCCGCCAGTGCCCAAGAT-----210
 QY 107 AlaGlnIleArgMetIleGlnIleSerLeuArgPheAlaSerValSerGluArgSer 126
 DB 211 -----TTCATCCCAATCGAATAATTGATTGTGC-----CATGCCATGGCC 249
 QY 127 LysTrp-----ProLysSerIleValAspValGlyCysGly 138
 DB 250 CAGTGGGGCGGATTAGATACACTTCCCGCGCAACAGGTATGATGGGTGGCGGC 309
 QY 139 IleGlyGlySerSerArgTrpLeuAlaLysPheGlyAlaThrSerValGlyIleThr 158
 DB 310 ATTGGCGGAGACGATGCTTCTCGCCAAAGATATTATGTTTAACGTTACCGGCGATCAC 369

QY 159 LeuSerProValGlnAlaGlnArgAlaAsnAlaLeuAlaIleAlaGlnGlyLeuAlaAsp 178
 DB 370 ATTACTCCCAACAGGTAACAGGGGCGACGAATTACTCTCCCGATGTACGGCC---426
 QY 179 LysValSerPheGlnIleAlaAlaAspAlaLeuGlnInProPheSerAspGlyGlnPheAsp 198
 DB 427 -----AAGTTGGCGGTGAGCATGTATGCTTGTCTTTCTTCGACGAGTATTGAC 480
 QY 199 LeuValTrpSerMetGluSerGlyGlnHisMetProAspLysAlaLysPheValGlyLeu 218
 DB 481 GTAGTTGTGGTGGTGGAGGAGGCGCCCATGCTGACAAAGCTGTGTTCGCAAGGAA 540
 QY 219 LeuAlaArgValAlaAlaProGlyValIleIleIleIleValIleThrTrpCysHisArgAsp 238
 DB 541 TTACTCGGGGTGTAACACAGGGGCGATTTGTTGGTGGCGGATTGGAATACAGCGGAC 600
 QY 239 LeuGlyProAspGluGln---SerLeuHisProTrpGluGlnAspLeuLysLysIle 257
 DB 601 -----GATCGCCAGGTGCCCTCAACTTCTGGAAAAACAGTATGCGACACTG 651
 QY 258 CysAspAlaTrpTrpLeuProAlaTrpCysSerThrSerAspTrpValLysLeuLeuGln 277
 DB 652 TTGATCAATGCTCCACCTGCTTTCACAGCATGAGGTTTTCGCGAAATTTGAA 711
 QY 278 SerLeuSerLeuGlnAsp-----IleLysSerGluAspTrpSerArgPheValAlaPro 295
 DB 712 GCCACGGGTTGGTGGAGGCCAGGAGTACTGCTGATGAGACTGTACCGACCTCC 771
 QY 296 PheTrpProAlaValIleArgSerAlaPheThrTrpLysGlyLeu 310
 DB 772 GCTTGGTTGGATACCATT-----TGGCAGGGCATT 801

Search completed: March 19, 2003, 08:05:54
 Job time : 2726 secs

FT /product- gamma-tocopherol_methyltransferase
 XX
 PN W0200032757-A2.
 XX
 PD 08-JUN-2000.
 XX
 PF 02-DEC-1999; 99WO-US28588.
 XX
 PR 03-DEC-1998; 98US-0110781.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E.I.
 XX
 PI Cahoon RE, Coughlan SJ, Miao G, Rafalski JA;
 XX
 DR WPT: 2000-412309/35.
 XX
 DR P-PSDB: AAY96472.
 XX
 PT Polynucleotide encoding a vitamin E biosynthetic enzyme polypeptide
 PT such as gamma tocopherol methyltransferase, useful for synthesizing
 PT vitamin E or as a target for design and discovery of herbicides
 XX
 PS Claim 3; Page 64-65; 82pp: English.
 XX
 CC AAA29151-69 encode vitamin E (alpha-tocopherol) biosynthetic enzymes or
 CC their fragments. The cDNA clones were identified by BLAST searches based
 CC on similarity to Synecocystis sp. and Arabidopsis thaliana vitamin E
 CC biosynthetic enzymes. The enzymes are useful for synthesizing
 CC plastoquinones or tocopherols (especially vitamin E). Vitamin E is
 CC required for plant growth, therefore the enzymes may be used for the
 CC discovery of new herbicides. The enzymes can be used in methods to
 CC evaluate potential inhibitors, which may have use as herbicides.
 CC Additionally, transgenic expression of, e.g. gamma-tocopherol
 CC methyltransferase affords the ability to manipulate tocopherol levels as
 CC desired for a particular application.
 XX
 SQ Sequence 1189 BP; 333 A; 253 C; 299 G; 304 T; 0 other:
 XX
 Alignment Scores:
 Pred. No.: 2,34e-180 Length: 1189
 Score: 1830.00 Matches: 350
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0
 US-09-857-613a-28 (1-350) x AAA29164 (1-1189)
 QY 1 MetalathrValAlaArgIleProThrIleSerCysIleHisIleHisThrPheArgSer 20
 DB 5 ATGGCCACCGGTGGAGATGCCAACAATCATGCATGCCATCCACACGCTTCCGTTCC 64
 QY 21 GluSerProArgThrPheAlaArgIleArgValGlyProArgSerTTPAlaProIleArg 40
 DB 65 CAATCCCTCCGACATTTCCGCGAATCCGCGTCCGACCCACAGCTGTGGCTCTATTCCG 124
 QY 41 AlaSerIleAlaIleSerSerGluArgGlyGluIleValleuGluGlnIlyProLysLysasp 60
 DB 125 GCATCGGCGAGGAGCTGGAGAGGAGGAGATGATTGGAGCGCAACCCCAAGAGGAT 184
 QY 61 AsplyLysLysLeuGlnLysGlyIleAlaGluPheTyrAspGluSerSerGlyLeuTyrP 80
 DB 185 GACAGAAGAAGACGCGAAGAGGAAATCGCAAGATTTCACGACGATCTCTGCGTTATGG 244
 QY 81 GluAsnIleTTPGlyAspHisMetHisIleGlyPheTyrAspSerAspSerThrValSer 100
 DB 245 GAGAACATTTTGGGCGACACATGACCATGCTTTATGACTGCGATTCACATGTTTCG 304
 QY 101 LeuSerAspHisArgAlaAlaGlnIleArgMetIleGlnIlySerIleArgPheAlaSer 120
 DB 305 CTTTCGATCATCTGCTGCTGCTCAATCGAATGCCAAGATCTCTTCGCTTTCCTCT 364
 QY 121 ValSerGluGluArgSerLysTTPProLysSerIleValAlaValGlyCysGlyIleGly 140

DB 365 GTTTCGAGAGCGCTAGTAATAGGCCCAAGAGTATAGTTGATGTGGGTGCGATAGT 424
 QY 141 GlySerSerArgTyrIleAlaIleLysPheGlyAlaThrSerValGlyIleThrLeuSer 160
 DB 425 GGCAGCTCTGATACCTGGCCAAAGAAATTTGGACCAACACATGATGCGATCTTGAGT 484
 QY 161 ProValGlnAlaGlnArgAlaAsnAlaLeuAlaIleAlaGlnIlyLeuAlaAspLysVal 180
 DB 485 CTTGTTCAAGCTCAAGAGCAAAATGCTCTCTGCTGCTCAAGATTTGGCTGATAGGTT 544
 QY 181 SerPheGlnValAlaAspAlaLeuGlnIleProPheSerAspGlyGlnPheAspLeuVal 200
 DB 545 TCTTTTCAAGTTGCTGAGAGCTCTACAGAACCATTTCTTGACGGCGCATTTGATCTGCTG 604
 QY 201 TrpSerMetGluSerGlyGlnHisMetProAspLysAlaLysPheValGlyLeuAla 220
 DB 605 TGGTCCATGAGAGATGGAGACATATGCTTCAAAAGCTTAAGTTGTTGGAGAGTTACT 664
 QY 221 ArgValAlaAlaProGlyAlaIleIleIleIleValThrTTPCysHisArgAspLeuGly 240
 DB 665 CGGTTAGCAGCACCGAGTGCCATTATATATAGTACATGCTGCGCACAGGAGATCTTGGC 724
 QY 241 ProAspGluGlnSerLeuHisProTTPGluGlnAspLeuLysLysIleCysAspAla 260
 DB 725 CCTGACGCAACAATCCTTACATCCATGGAGCAAGATCTTAAAGAAATTTGCGATGCA 784
 QY 261 TyrTyrIleProAlaIleTTPCysSerThrSerAspLysValLysLeuGlnIleSerLeuSer 280
 DB 785 TATTACTCTCCCTGCTGCTCAACTTGAATATGTTAAGTTGCTCAATCCCTGTCGA 844
 QY 281 LeuGlnAspIleLysSerGluAspTTPSerArgPheValAlaProPheTTPProAlaVal 300
 DB 845 CTTAGACATCAAGATCGAAGAGATTTGCTGCTTGTGCTCCATTTTGGCCACAGCTG 904
 QY 301 IleArgSerAlaPheThrTTPLysGlyLeuSerSerLeuLeuSerSerGlyGlnLysThr 320
 DB 905 ATACGCTACGCTTCACATGAGAGGCTATCTTCACTTGGAGACAGCAAAAGAG 964
 QY 321 IleLysGlyAlaLeuAlaMetProLeuMetIleGluGlyTyrLysLysAspLeuIleLys 340
 DB 965 ATAAAGAGAGCTTGGCTATGCTCATTTGATGATAGAGGATACAGAAAGATCTAATTAAAG 1024
 QY 341 PheAlaIleIleThrCysArgLysProGlu 350
 DB 1025 TTTGCCATCATTCATGTCGAAACCTGAA 1054
 RESULT 2
 AAA29154
 ID AAA29154 standard; cDNA; 1331 BP.
 XX
 AC AAA29154;
 XX
 DT 12-SEP-2000 (first entry)
 XX
 DE Soybean gamma tocopherol methyltransferase contig cDNA.
 XX
 KW Vitamin E: alpha-tocopherol; biosynthesis; enzyme; inhibitor; herbicide;
 KW gamma-tocopherol methyltransferase; ss.
 XX
 OS Glycine max.
 XX
 FH Key Location/Qualifiers
 FT CDS 5..1055
 FT /*tag= a
 FT /product= gamma_tocopherol_methyltransferase
 PN W0200032757-A2.
 XX
 PD 08-JUN-2000.
 XX
 PF 02-DEC-1999; 99WO-US28588.
 XX
 PR 03-DEC-1998; 98US-0110781.

XX (DUPO) DU PONT DE NEMOURS & CO E I.
 PA
 XX
 XX Cahoan RE, Coughlan SJ, Miao G, Rafalski JA:
 XX
 DR WPI: 2000-412309/35.
 XX P-PSDB: AAY96462.
 XX
 PT Polynucleotide encoding a vitamin E biosynthetic enzyme polypeptide
 PT as gamma tocopherol methyltransferase, useful for synthesizing
 PT vitamin E or as a target for design and discovery of herbicides
 XX
 XX
 PS Claim 3: Page 52: 82pp: English.
 XX
 XX AAA29151-69 encode vitamin E (alpha-tocopherol) biosynthetic enzymes or
 CC their fragments. The cDNA clones were identified by BLAST searches based
 CC on similarity to Synecocystis sp. and Arabidopsis thaliana vitamin E
 CC biosynthetic enzymes. The enzymes are useful for synthesizing
 CC plastiquinones or tocopherols (especially vitamin E). Vitamin E is
 CC required for plant growth, therefore the enzymes may be used for the
 CC discovery of new herbicides. The enzymes can be used in methods to
 CC evaluate potential inhibitors, which may have use as herbicides.
 CC Additionally, transgene expression of, e.g. gamma-tocopherol
 CC methyltransferase affords the ability to manipulate tocopherol levels as
 CC desired for a particular application.
 CC
 XX
 XX Sequence 1331 BP: 373 A; 293 C; 301 G; 364 T; 0 other:
 SQ
 Alignment Scores:
 Pred. No.: 1,666-163 Length: 1331
 Score: 1668.50 Matches: 323
 Percent Similarity: 94.52% Conservative: 5
 Best Local Similarity: 93.08% Mismatches: 10
 Query Match: 91.17% Indels: 9
 DB: 21 Gaps: 1
 US-09-857-613a-28 (1-350) x AAA29154 (1-1331)
 QY 1 MetAlaThrValAlaArgIleProThrIleSerCysIleHisIleHisThrPheArgSer 20
 DB 6 ATGGCCACCGTGGTGAAGATCCCAACATCTCATGTCACATCCACAGCTCCGCTCC 65
 QY 21 GlnSerProArgThrPheAlaArgIleArgValGlyProArgSerTrpAlaProIleArg 40
 DB 66 CAATCCCTCCGCACTTCCGCCAATCCGGGTGGAGCCAGGCTGGCTCTATTCCG 125
 QY 41 AlaSerAlaIleSerSerGluArgGlyIleValIleGluGlnIleProLysAsp 60
 DB 126 GCATTCGGACGCTCGAGAGAGGGGAGATGATGAGGAGCAAGCCGAGAGAGAT 185
 QY 61 AsplysLysLysLeuGlnIleGlyIleAlaGluPheTrpAspGlnSerSerGlyLeuTrp 80
 DB 186 GACAAAGAGAGCTGCAGAGAGGAGATCGCAGATTTTACGACAGTCGTCGCTATG 245
 QY 81 GluAsnIleTrpClyAspHisMetHisIleGlyPheTrpAspSerSerTrpValSer 100
 DB 246 GACAAATTTGGGGCCACACATGACATGCTTTATGATCGATTCACATGTTTCG 305
 QY 101 LeuSerAspHisArgAlaAlaGlnIleArgMetIleGlnGlnSerLeuArgPheAlaSer 120
 DB 306 CTTTCGATCATCGCTCGCTCGATCCGATATCCAAAGAGCTCTTCGCTTGCCTCT 365
 QY 121 ValSerGlnGluArgSerLeuTrpProLysSerIleValAspValGlyCysGlyIleGly 140
 DB 366 GTTTCGAGAGAGCTGAATATGCCCCAGAGATATGTTGCTGGCTGGCATAGAT 425
 QY 141 GlySerSerArgTrpLeuAlaLysLysPheGlyAlaThrSerValGlyIleTrpLeuSer 160
 DB 426 GCGAGCTCTAGATACCTGGCCAGAAATTTGGAGCAACAGTAGGACATCAGCTAGT 485
 QY 161 ProValGlnAlaGlnArgAlaAsnAlaLeuAlaAlaIleGlnGlyLeuAlaAspLysVal 180
 DB 486 CCGTTCAACCTCAAAAGAGCAAAATGCTTTCGCTGCTCAAGGATTGGCTGATTAAGTT 545

QY 181 SerPheGlnValAlaAspAlaLeuGlnIleProPheSerAspGlyGlnPheAspLeuVal 200
 DB 546 TCCTTTCAGGTTGCTGACGCTCTACAGCAACCATCTCTGACGGCCATTTGATCTGGTG 605
 QY 201 TrpSerMetGluSerGlyGluHisMetProAspLysAlaLysPheValGlyGluLeuAla 220
 DB 606 TGGTCATGAGACAGTGCAGAGCATATGCTGCACAAAGCTAAGTTGTTGGAGAGTTACT 665
 QY 221 ArgValAlaAlaProGlyAlaIleIleIleValIleThrProCysHisArgAspLeuGly 240
 DB 666 CGGGTAGACAGCACAGGTGCCATTAATAATAGTAACATGGTGCACAGGATCTTGCC 725
 QY 241 ProAspGlnGlnSerLeuHisProTrpGlnGlnAspLeuLeuLysLysIleCysAspAla 260
 DB 726 CCGAGCAACATCTTACATCCATGCGGAGCAAGATCTTTAAAGAGATTGTGCAATGCA 785
 QY 261 TyrTrpLeuProAlaIleTrpCysSerThrSerAspTrpValLysLeuGlnSerLeuSer 280
 DB 786 TATTACCTCCCTGCTGCTGCTCAACTTCTGATTTATGTTAAGTTGCTCCAAATCCCTGCA 845
 QY 281 LeuGlnAspIleLysSerGluAspTrpSerArgPheValAlaProPheTrpProAlaVal 300
 DB 846 CTTCAAGACATCAAGTACAGACATTTGCTCGCTTGTGCTGCATTTTGGCCACAGTGA 905
 QY 301 IleArgSerAlaPheThrTrpLysGlyLeuSerSerLeuLeuSerSerGlyGlnLysThr 320
 DB 906 ATACGCTCAGCCTTCACATGGAAGGCTTATCTTACCTTGAGCAGTGTAAGCTTGCA 965
 QY 321 IleLysGlyAlaLeuAlaMetProLeuMetIleGlyIleTyrLysLysAspLeuIleLys 340
 DB 966 ATTTATATTGCA-----TTTCAAAACAAACCCCCCA 998
 QY 341 PheAlaIleIleThrCysArg 347
 DB 999 TCTTATTGCAACTTGCAAG 1019
 RESULT 3
 AAA29165
 ID AAA29165 standard; cDNA; 1257 BP.
 XX
 AC AAA29165;
 XX
 DT 12-SEP-2000 (first entry)
 XX
 DE Soybean gamma-tocopherol methyltransferase cDNA.
 XX
 KW Vitamin E: alpha-tocopherol; biosynthesis; enzyme; inhibitor; herbicide;
 KW gamma-tocopherol methyltransferase; ss.
 XX
 OS Trilicium aestivum.
 XX
 FH Key Location/Qualifiers
 FT CDS 18..1118
 FT /tag= a
 FT /transl_except= (pos:30..32, aa:Xaa)
 FT /transl_except= (pos:150..152, aa:Xaa)
 FT /product= gamma-tocopherol methyltransferase
 FT /note= "Xaa is not defined"
 PN
 XX
 XX WO200032757-A2.
 PD 08-JUN-2000.
 XX
 PE 02-DEC-1999: 99NO-US28588.
 XX
 PR 03-DEC-1998: 98US-0110781.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 XX Cahoan RE, Coughlan SJ, Miao G, Rafalski JA:
 DR WPI: 2000-412309/35.

DR P-PSDB: AAY96473.

PT Polynucleotide encoding a vitamin E biosynthetic enzyme polypeptide
 PT such as gamma tocopherol methyltransferase, useful for synthesizing
 PT vitamin E or as a target for design and discovery of herbicides

PS Claim 3; Page 66; 82pp; English.

CC AAA29151-69 encode vitamin E (alpha-tocopherol) biosynthetic enzymes or
 CC their fragments. The cDNA clones were identified by BLAST searches based
 CC on similarity to *Synechocystis* sp. and *Arabidopsis thaliana* vitamin E
 CC biosynthetic enzymes. The enzymes are useful for synthesizing
 CC plastoquinones or tocopherols (especially vitamin E). Vitamin E is
 CC required for plant growth, therefore the enzymes may be used for the
 CC discovery of new herbicides. The enzymes can be used in methods to
 CC evaluate potential inhibitors, which may have use as herbicides.
 CC Additionally, transgenic expression of, e.g. gamma-tocopherol
 CC methyltransferase affords the ability to manipulate tocopherol levels as
 CC desired for a particular application.

SQ Sequence 1257 BP; 299 A; 370 C; 343 G; 243 T; 2 other:

Alignment Scores:

Pred. No.:	4,76e-112	Length:	1257
Score:	1173.00	Matches:	224
Percent Similarity:	80.828	Conservative:	33
Best Local Similarity:	70.448	Mismatches:	61
Query Match:	64.10%	Indels:	0
DB:	21	Gaps:	0

US-09-857-613A-28 (1-350) x AAA29165 (1-1257)

OY 33 ProArgSerTrpAlaProIleArgAlaSerAlaAlaSerSerIuArgIleVal 52
 DB 117 CCGCGCCTGTCTCTCCCTCCGCCGAGACGCTCCGTCGCCCGATGCGCTC 176
 OY 53 LeuGluGlnLysProLysLysAspAspLysLysLysLysGlnLysGlnLysGlnLys 72
 DB 177 GAGGAGCTC 236
 OY 73 TyrAspLysSerSerLysLysLysLysLysLysLysLysLysLysLysLysLys 92
 DB 237 TAGCAGAGAGTCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 296
 OY 93 TyrAspSerAspSerThrValSerLysSerAspHisArgAlaAlaGlnIleArgMetIle 112
 DB 297 TAGCAGTCGATC 356
 OY 113 GlnGluSerLeuArgPheAlaSerValSerGlnGluArgSerLysTrpProLysSerIle 132
 DB 357 GAGGAGCGATT 416
 OY 133 ValAspValGlyCysGlyIleGlyLysSerSerArgTyrLeuAlaLysLysPheGlyAla 152
 DB 417 GTTGATGTTGATGATCGGAAATCGGTGAGCTCAAGATACCTGCGCAACAATATGAGACA 476
 OY 153 ThrSerValGlyLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 172
 DB 477 CAATGCTCTGAGTACATGAGTACGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 536
 OY 173 AlaGlnGlyLeuAlaLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 192
 DB 537 GCGCAGGCGGTTGTCGAGCAAGGCTTCTTCCAGTTCCTATGCTGAGCAACCATTTT 596
 OY 193 SerAspGlyGlnPheAspLeuValTrpSerMetCysSerLysLysLysLysLysLys 212
 DB 597 CCGATGAGCGAGTGTATGATCTTGTCTGTATGAGAGTGTGTAGCAGCAGCCGAAACAA 656
 OY 213 AlaLysPheValGlyLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 232
 DB 657 CAGAAAGTTTGAAGCAGAGTCGACCGCGTCCAGAGCTCCAGAGAGAACTATCATCATGCTG 716
 OY 233 ThrTrpCysHisArgAspLeuGlyProAspGlnGlnSerLeuHisProTrpGlnGlnAsp 252

DB 717 ACCTGTCGATAGAGAACTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 776
 OY 253 LeuLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 772
 DB 777 CTTTGAAGAAATTTGATGATCATATACCTCCCGATGAGTGTGCGCGCGCGCGCGCGCT 836
 OY 273 ValLysLeuLeuGlnSerLeuGlnAspIleLysSerGlnLysProLysSerArgPhe 292
 DB 837 GTCMAATTTGCGAGCATGCTCTTGAAGATATCAAAACGCGCGCGCGCGCGCGCGCG 896
 OY 293 ValAlaProPheTrpProAlaValIleArgSerAlaPheTrpLysLysLysSerSer 312
 DB 897 GTGCGCGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 956
 OY 313 LeuLeuSerSerLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 332
 DB 957 CTACTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1016
 OY 333 GlyTyrLysLysAspLeuLysLysLysLysLysLysLysLysLysLysLysLysLys 350
 DB 1017 GGCTACAGAAAGGCTCATTAAGTTCACATCATCCTGCGCGCGCGCGCGCGCGCG 1070

RESULT 4
 ID AAA29161 standard; cDNA: 1102 BP.
 XX AAA29161;
 AC 12-SEP-2000 (first entry)
 DT
 XX
 DE Corn gamma-tocopherol methyltransferase cDNA.
 KW Vitamin E; alpha-tocopherol; biosynthesis; enzyme; inhibitor; herbicide;
 KW gamma-tocopherol methyltransferase; ss.
 XX
 OS Zea mays.
 XX
 FH
 FT
 FT CDS
 FT Location/Qualifiers
 FT 1..1059
 FT /*tag= a
 FT /product= gamma-tocopherol methyltransferase
 FT /transl_except= (pos:454..456, aa:xaa)
 FT /note= "xaa not defined"

WO200032757-A2.
 PD 08-JUN-2000.
 PD
 PF 02-DEC-1999: 99WO-US28588.
 XX
 PR 03-DEC-1998: 98US-0110781.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 XX
 PI Cahoon RE, Coughlan SJ, Miao G, Rafalski JA;
 XX
 XX WPI: 2000-412309/35.
 DR P-PSDB: AAY96469.
 PT Polynucleotide encoding a vitamin E biosynthetic enzyme polypeptide
 PT such as gamma tocopherol methyltransferase, useful for synthesizing
 PT vitamin E or as a target for design and discovery of herbicides

PS Claim 3; Page 60; 82pp; English.

CC AAA29151-69 encode vitamin E (alpha-tocopherol) biosynthetic enzymes or
 CC their fragments. The cDNA clones were identified by BLAST searches based
 CC on similarity to *Synechocystis* sp. and *Arabidopsis thaliana* vitamin E
 CC biosynthetic enzymes. The enzymes are useful for synthesizing
 CC plastoquinones or tocopherols (especially vitamin E). Vitamin E is
 CC required for plant growth, therefore the enzymes may be used for the
 CC discovery of new herbicides. The enzymes can be used in methods to

CC evaluate potential inhibitors, which may have use as herbicides.
CC Additionally, transgenic expression of, e.g. gamma-tocopherol
CC methyltransferase affords the ability to manipulate tocopherol levels as
CC desired for a particular application.

50 Sequence 1102 BP; 231 A; 330 C; 328 G; 207 T; 6 other:

Alignment Scores:

Pred. No.:	5.5e-111	Length:	1102
Score:	1162.00	Matches:	225
Percent Similarity:	75.88%	Conservative:	33
Best Local Similarity:	66.18%	Mismatches:	72
Query Match:	63.50%	Indels:	10
DB:	21	Gaps:	2

US-09-857-613a-28 (1-350) x AAA29161 (1-1102)

```
OY 20 SerGlnSerProArgThrPheAlaArgValGlyProArgSerTrpAlaProIle 39
    |||||
DB 28 TCCTGAGCTCCAGGAGCTCCGACCTGCGGCGCGGAGCAGTACCTACCGCCCTTCG 87
OY 40 ArgAlaSerAlaAlaSerSerGluArgGlyGluIleValLeuGlnGlnLysProLys 59
    |||
DB 88 CACGTCCCGCGGCACTCCCGCGCTCCGACGCGCGCTGTCAGCCTGGCGATGGCC 147
OY 60 AspAspLysLysLys-----LeuGlnLysLysLysLysLysLysLysLysLys 71
    :
DB 148 TCGTCGACGCGCTCAGCGCCCGCGCGCGCGCGCTCTGTAAGGAGGCGATCGCGGG 207
OY 72 PheTrpAspGlnSerSerGlyLeuTrpGluAsnIleTrpGluAlaPheIleMet 91
    |||||
DB 208 CTGTACGAGCTGCTCGGCGCTGCGGAGAACATCTGGGCGACCATGACACCGCGG 267
OY 92 PheTrpAspSerAspSerThrValSerLeuSerAspHisArgAlaAlaGlnIleArgMet 111
    |||||
DB 268 TTCCTACGACTCGAGGAGGAGCCCTCCATCGGCGATCACCGCGCGCGCGATCCGATG 327
OY 112 TleGlnGlnSerLeuArgPheAlaSerVal-----SerGlnGlnArgSerLysTrpPro 129
    |||||
DB 328 ATCGAGGAGGCGCTCGCTCCGCGCTGTCAGCGCTGTCAGCATGATCAGAGAACACCA 387
OY 130 LysSerIleValAlaGlnGlyCysGlyIleGlySerSerArgTrpLeuAlaLys 149
    |||||
DB 388 AAACATATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATG 447
OY 150 PheGlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnAlaAla 169
    :
DB 448 TACGAGAACCGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTG 507
OY 170 LeuAlaAlaAlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAlaLeuGln 189
    |||||
DB 508 CTCGCTGACGCGAGGAGGTTGTCGATCAGTACTCTCAAGTCTGATGCTCGAG 567
OY 190 GlnProPheSerAspGlyLysIlePheAspLeuValTrpSerMetLysSerGlyLysMet 209
    |||||
DB 568 CAACGTTTCTTCCAGCGGAGTTCGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 627
OY 210 ProAspLysAlaLysPheValGlyGlnLeuAlaArgValAlaAlaProGlyAlaIle 229
    |||||
DB 628 CCGGACAGAGAAAGTTTCTTCTAGTACGACGCGCTGCGGCTCTCTGAGGAGGACATA 687
OY 230 IleIleValThrTrpCysHisArgAspLeuGlyProAspGlnGlnSerLeuHisProTrp 249
    |||||
DB 688 ATCATGCTGACATGTCGATGAGAACGTGATCCGAAACCTCGCTAAAGCCGAT 747
OY 250 GluGlnAspLeuLysLysLysLysCysAspAlaTrpTrpLeuProAlaTrpCysSerThr 269
    |||||
DB 748 GAACCTGACCTCTCTGAGGAGATATCCAGCGCTACTACCTCCGAGCTGCTCCCT 807
OY 270 SerAspTrpValLysLeuGlnSerLeuSerLeuGlnAspIleLysSerGlnAspTrp 289
    |||||
DB 808 TCAGATATGTGAAATTCGCAAGTCACTCTCTCTGAGAGATATCAAGACAGCTGACTGG 867
```

```
OY 290 SerArgPheValAlaProPheTrpProAlaValIleArgSerAlaPheThrTrpLysGly 309
    |||
DB 868 TCAGGAGAACGTGCGCCGCTTTGGCCCGCTGATTAATTCAGCCCTAACATGCAAGGCG 927
OY 310 LeuSerSerLeuLeuSerSerGlyGlnLysThrIleLysGlyAlaLeuAlaMetProLeu 329
    :
DB 928 TTCACCTCTCTCTGACGACCGGATGGAAGACGATCAGAGCGCGCATGTGATGCCGCTA 987
OY 330 MetIleGlnGlyTrpLysLysAspLeuIleLysPheAlaIleIleThrCysArgLysPro 349
    |||||
DB 988 ATGATTCAGGCTACAGAAAGGCGCTCATCAATTCACCATCATCAGCTGTGCAAGCCT 1047
RESULT 5
AAC50535
ID AAC50535 standard; DNA: 1293 BP.
AC AAC50535;
DE 18-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 65205.
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 65205.
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
```

```
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154479.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 4,15e-110
Score: 1154.50
Percent Similarity: 75.43%

Length: 1293
Matches: 231
Conservative: 30
```

Best Local Similarity:	66.76%	Mismatches:	65
Query Match:	63.09%	Indels:	21
DB:	21	Gaps:	4

US-09-857-613A-28 (1-350) x AAC50535 (1-1293)

Oy	8	ProthrlleSerCysIleHisIleThrPheArgSerGlnSerProArgThrPheAla	27
Db	118	CGAACTCTTCTTCCGGCTCAAAAGTCATCCGCTCTTT-CGGTCTCCATCTCTCTCC	176
Oy	28	ArgIleArgVal---GlyProArgSerTrpAlaPolleArgAlaSerAlaIleSer	46
Db	177	TCAGTCTCATGACGACAACCGGTGGAAACGTCGTGGCGGCTGCTACATCCACT	236
Oy	47	GluArgGlyGluIleValLeuGlnGlnLysProLysLysAspLysLysLysLeuGln	66
Db	237	GAG-----GCGCTAGA	248
Oy	67	LysGlyIleAlaGluPheTyrAspGlnSerSerGlyLeuTrpGluAsnIleTrpGlyAsp	86
Db	249	AAAGGAATACCGGAGATTCATCAATGAACCTTCGGGTTGTGGGAAGAGATTGGGGAGAT	308
Oy	87	HisMetHisHisGlyPheTyrAspSerAspSerThrValSerLeuSerAsp-----His	104
Db	309	CATTCGCATCATGCGCTTTATGACCCCTGACCTTCCTCTGTCACACTTTCGATTCTGCTAC	368
Oy	105	ArgAlaIleAlaGlnIleArgMetIleGlnGlnSerLeuArgPheAlaSerValSer---Glu	123
Db	369	AAGAAACCTCAGATCCGTATGATGAAGACTCTCCGCTTGGCCGCTTACGATGA	428
Oy	124	GluArgSerLysTrpProLysSerIleValAlaSpValGlyCysGlyIleGlyLysSer	143
Db	429	GAGAGAGAGAAAAGATTAAGAAGATGATGATGTTGGTGCGGATTCGAGGAAGCTCA	488
Oy	144	ArgTyrLeuAlaLysLysPheGlyAlaThrSerValGlyIleThrLeuSerProValGln	163
Db	489	AGATATCTTCCCTCTAAATTTGGAGCTGATCATCGCATACCTCAAGCCCTGTTCAG	548
Oy	164	AlaGlnArgAlaAsnAlaLeuAlaAlaGlnGlyLeuAlaAspLysValSerPheGln	183
Db	549	GCCAAAGAGACCATGATCTCCGGGCTGCTCATCTACCTGCCATAAAGCTTCCTTCCAA	608
Oy	184	ValAlaAspAlaLeuGlnGlnIleProPheSerAspGlyGlnPheAspLeuValTrpSerMet	203
Db	609	GTTCCGATCGCTGGATCGATCAGCATTCGAAGATGGAATTCGATCTAGTGGGTCATG	668
Oy	204	GluSerGlyLuhHisMetProAspLysAlaLysPheValGlyGluLeuAlaArgValAla	223
Db	669	GAGAGTGTGAGCATATGCTCTGACAAGGCCAAGTTGTCTAAAGAGCTGTGCTGCTGCG	728
Oy	224	AlaProGlyAlaIleIleIleIleValIleThrProCysHisArgAspLeuGlyProAspGlu	243
Db	729	GCTCCAGAGAGTGAAGATAATAGTACATGGTGGCATAGAAATCTATCTGCGGGGAG	788
Oy	244	GlnSerLeuHisProTrpGlnGlnAspLeuLysLysIleCysAspAlaThrIleu	263
Db	789	GAACTTTTGACGCGGTGGAGCAAAACATCTTGAGCAAAATCTGTAAAGACGTTTATCTC	848
Oy	264	ProAlaTrpCysSerThrSerAspTyrValLysLeuLeuGlnSerLeuSerLeuGlnasp	283
Db	849	CCGCGCTGGTCTCCACCCGATGATATATGTCACATTCCTTCATCCCATCTCTCCAGAT	908
Oy	284	IleLysSerGluAspTrpSerArgPheValAlaProPheTrpProAlaValIleArgSer	303
Db	909	ATTAAAGTCTGGGATTTGGTACAGAAACGTAAGCTCTTTCTGGGCTCCGTTATACGACT	968
Oy	304	AlaPheThrTrpLysGlyLeuSerSerLeuLeuSerSerGlyGlnLysThrIleLysGly	323
Db	969	GCATTAAACATGGAAGGCGCTTGCTCTCTCTCCGTAAGCGTATGAAGAAATATTAAGCA	1028
Oy	324	AlaLeuAlaMetProLeuMetIleGlnGlyTyrLysLysAspLeuIleLysPheAlaIle	343
Db	1029	GCATTGACAATGCCATTGATGATTTGAAGGTTCACAAAGAGCTCACTAATGATTTGCTATC	1088

QY	344	11eThrcysarglyspro	349
		:::	
Db	1089	ATCACTTGCCAGAAGCCA	110

RESULT 6

ID	AAZ98317	standard; DNA; 1071 BP
xx		

AC	AAZ98317
XY	

DT 14-JUN-2000 (first entry)
XX

DE A. Chaliana gene: involved in environmental stress tolerance

KM environmental stress; plant; anaerobic; flooding, cold
dehydration; drought; heat stress; salinity; osmotolerance; ds.

Arabidopsis thaliana.

PN WO200008187-A2.

PD 17-FEB-2000

PF 04-AUG-1999; 99WO-EP05652.

PR 04-AUG-1998; 98EP-0202634.

PA (VLA-) VLAMS INTERUNIVERSITAIR INST BIOTECHNOG

PI Lee JH, Verbruggen N;
v

DR WPI; 2000-205726/18.
DP P-PSDB: AAV77937

XX Isolation of poly

XX p1 Isolating genes involved in tolerance to environmental stress -

PS CLAIM 4; Page 130-132; 312pp; English
XX

CC The invention relates to isolation of coding sequences and/or genes
CC involved in tolerance to environmental stress in plants. The sequences
CC (AA298305-298365) are useful for producing a transgenic plant having
CC enhanced tolerance or resistance to environmental stress conditions such
CC as anaerobic, flooding, cold, dehydration, drought, heat stress or
CC salinity. This is useful for producing improved yield, growth,
CC development and productivity under environmental stress conditions, and
CC also provides growth of crops in areas where they cannot grow without
CC the induced osmocompatibility. Sequences AA298305-365 represent
CC polynucleotide sequences from *A. thaliana* that are involved in
CC environmental stress tolerance.

Sequence 1071 BP; 277 A; 229 C; 270 G; 295 T; 0 other;

Alignment Scores:

Score:
Percent Similarity

Query Match:	62.92%	Indels:	21
Best Local Similarity:	66.47%	Mismatches:	63

DB:	21	caps:	4
-----	----	-------	---

US-09-03/013A-Z6 (1-320) X A442001/ (1-10/1)

[illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

[illegible][illegible]

```

180 GAG-----GCGGTAGA 191
QY 67 LysGlyIleAlaGluPheTyrAspGluSerSerGlyLeuTrpGluAlaIleTrpGlyAsp 86
192 AAGAGTAATAGCGAGTCTTACATGAATGAACCTTCGGTTGTGGAGACATTTGGGAGAT 251
QY 87 HisMetHisHisGlyPheTyrAspSerSerPheThrValSerLeuSerAsp-----His 104
252 CATATGCAATCATGCTTTATGACCCCTGATCTTCTGTTCAACTTTCGATTCGTGTGAC 311
QY 105 ArgAlaIleAlaGluIleArgMetIleGlnGluSerLeuArgPheAlaSerValSer---Glu 123
312 AAGGAAGCTCATGATCCGATCATTAATGAAGCTCTCCGTTCCGGGCTTACTATGA 371
QY 124 GluArgSerLysTrpProLysSerIleValAspValGlyCysGlyIleGlyLysSer 143
372 GAGGAGGAGAAAAAGATAAAGAACTAGCTGATTTGGGTGGGATTTGGAGAGATCA 431
QY 144 ArgTyrLeuAlaLysLysPheGlyValThrSerValGlyIleThrLeuSerProValGln 163
432 AGATATCTTGGCTTAAATTTGGAGCTGAATGCAATGCAATTCATTCAGCCCTTTCAG 491
QY 164 AlaGlnArgAlaAsnAlaLeuAlaIleAlaGlnGlyLeuAlaAspLysValSerPheGln 183
492 GCCAAGACACCAATGATCTCGCGCTCTCATATCCTCTCATTAAGGCTTCTCCAA 551
QY 184 ValAlaAspAlaLeuGlnGlnProPheSerAspGlyGlnPheAspLeuValTrpSerMet 203
552 GTTGGGATGCGCTTGATGAGCCATTCGAGATCGAAATTCGATCTGATGTCGATG 611
QY 204 GluSerGlyGlnHisMetProAspLysAlaLysPheValGlyGlnLeuAlaArgValAla 223
612 GAGAGTGTGAGCAATATCTGACAGGCCCAAGTTGTAAAGAGTTGTAGCTGTGCGG 671
QY 224 AlaProGlyAlaIleIleIleValThrTrpCysHisArgAspLeuGlyProAspGlu 243
672 GCTCCAGAGAGTATGATATATATGATGATGATGATGATGATGATGATGATGATG 731
QY 244 GlnSerLeuHisProTrpGluGlnAspLeuLysLysIleCysAspAlaTrpTyrLeu 263
732 GAAGCTTTGACAGCCGTGGAGCAAAACATCTTGACAAATCTGTAGACGTTCTATCTC 791
QY 264 ProAlaTrpCysSerThrSerAspTyrValLysLeuGlnSerLeuSerLeuGlnAsp 283
792 CCGGCTGTGCTCCACCGATGATTAATGTCACCTTCCTCAATTCCTCTCCAGGAT 851
QY 284 IleLysSerGluAspTrpSerArgPheValAlaProPheTrpProAlaValIleArgSer 303
852 ATTAAGTGTGCGATTTGTCAGAGACGATGCTCTTCTGCGCTGCGGTTATAGGACT 911
QY 304 AlaPheTrpLysGlyLeuSerSerLeuLeuSerSerGlyGlnLysThrIleLysGly 323
912 GCATTAACATGAGAGGCGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 971
QY 324 AlaLeuAlaMetProLeuMetIleGluGlyTyrLysLysAspLeuIleLysHeAlaIle 343
972 GCATTGACAAATGCAATGATTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTA 1031
QY 344 IleThrCysArgLysPro 349
1032 ATCACTTGCCAGAACCA 1049

```

```

KW Vitamin E: homogenitase; HG: homogenitase-1,2-dioxygenase; HGD:
KW maleylacetoacetate-isomerase; MAI: fumarylacetoacetate-hydrolase; FAAH:
KW transgenic; nutrition; antioxidant; cardiovascular disease; cancer;
KW immune system; generalised age-related degeneration; animal feed;
KW meat quality; cosmetics; growth regulator; herbicide; cardiant;
KW cyclostatic; immunostimulant; enzyme; gene; ss.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
FT CDS 63..1109
FT /tag= a
FT /product= "gamma-tocopherol methyltransferase"
XX
PN WO200231173-A2.
XX
PD 18-APR-2002.
XX
PE 18-SEP-2001; 2001WO-EP10779.
XX
PR 19-SEP-2000; 2000DE-1046462.
XX
PA (SUNG-) SUNGENE GMBH & CO KGAA.
XX
PI Geiger M, Edmeth M, Kunze I:
XX
DR WPI: 2002-362682/39.
XX
DR P-PSDB: AA017426.
XX
PT Increasing Vitamin E production, useful in human or animal nutrition,
XX comprises reducing activity of enzymes that metabolize homogenitase
XX
PS Disclosure: Page 83-85; 98pp; German.
XX
CC The present invention relates to a method of producing Vitamin E,
CC involving modulating the synthesis of vitamin E by reducing decomposition
CC of homogenitase (HG) by reducing the activity of
CC homogenitase-1,2-dioxygenase (HGD), maleylacetoacetate-isomerase (MAI)
CC and/or fumarylacetoacetate-hydrolase (FAAH). The method is used to
CC produce transgenic organisms that are useful in animal and human
CC nutrition and for the isolation of vitamin E. Vitamin E is a fat-soluble
CC antioxidant with a protective effect against cardiovascular disease and
CC cancer. It also stimulates the immune system and may prevent generalised
CC age-related degeneration. When used in animal feeds, it improves quality
CC and storability of meat and can also be used in cosmetics. Also
CC antibodies raised against the 3 specified enzymes are useful in screening
CC for specific inhibitors, potentially useful as growth regulators, e.g.
CC herbicides. The present sequence is the A thaliana gamma-tocopherol
CC methyltransferase cDNA.
XX
SQ Sequence 1350 BP; 374 A; 278 C; 314 G; 384 T; 0 other:
XX
Alignment Scores:
Pred. No.: 9,04e-110 Length: 1350
Score: 1151.50 Matches: 230
Percent Similarity: 75.43% Conservative: 31
Best Local Similarity: 66.47% Mismatches: 65
Query Match: 62.92% Indels: 21
DB: 24 Gaps: 4
US-09-857-613A-28 (1-350) x AAL46039 (1-1350)
QY 8 ProPheIleSerCysIleHisIleHisThrPheArgSerGlnSerProArgThrPheAla 27
DB 115 CCACTCTCTCTTCGCGTCAAGTCAATGCTCTTTCGCTCCACACCTCCCTCCCTCC 173
QY 28 ArgIleArgVal---GlyProArgSerTrpAlaProIleArgAlaSerAlaIleSer 46
DB 174 TCAGTCTTATGACGACCAACGCGTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 233
QY 47 GluArgGlyGluIleValLeuGlnLysProLysLysAspLysLysLysLeuGln 66
DB 234 GAG-----GCGGTAGA 245

```

```

QY 67 LysGlyIleAlaGluPheTyrAspGluSerSerGlyLeuTrpGluAsnIleTrpGlyAsp 86
DB 246 AAGAGATAGCGGAGTTCACATGAACATCGCTTGGCAAGAGATTGGGAGAT 305
QY 87 HisMetHisHisGlyPheTyrAspSerAspSerThrValSerLeuSerAsp-----His 104
DB 306 CATATCATCATAGCTTTATATGACCTGATCTTCTGTTCACTTCTGATTTGCTGCTAC 365
QY 105 ArgAlaAlaGlnIleArgMetIleGlnGluSerLeuArgPheAlaSerValSer---Glu 123
DB 366 AAGGAGCTCAGATCCGCTATGATTGAAGACTCTCCGTTCCCGGCTGTACTGATTGA 425
QY 124 GluArgSerIleTrpProLysSerIleValAspValGlyCysGlyIleGlyGlySer 143
DB 426 GAGGAGCAGCAAAAAGATAAAGAAAGTACGATGCTGGCTGGATTTGAGGAGAGCTCA 485
QY 144 ArgTyrLeuAlaLysLysPheGlyAlaThrSerValGlyIlePheTrpLeuSerProValGln 163
DB 486 AGTATATCTTGCCTTAATTTGAGCTGAATTCGATTTGCTACTCTCAGCCCTGTTTACG 545
QY 164 AlaGlnArgAlaAsnAlaLeuAlaAlaAlaGlnGlyLeuAlaAspLysValSerPheGln 183
DB 546 GCCAAGAGACCAATGATATCGCGGCTGCTCATATCACTCTCATATAAGCGTCTCTCCAA 605
QY 184 ValAlaAspAlaLeuGlnGlnProPheSerAspGlyGlnPheAspLeuValTrpSerMet 203
DB 606 GTTCCGATCCGCTTGAATTCAGCATTCGAAAGTGAATAATTCATGATGCTGCTCATG 665
QY 204 GluSerGlyGluHisMetProAspLysAlaLysPheValGlyGluLeuAlaArgValAla 223
DB 666 GAGAGTGTGTAGCATATGCTGACAAAGGCCAAGTTGTATAAGGTTGGTACGTGTGGCG 725
QY 224 AlaProGlyAlaIleIleIleIleValThrTrpCysHisArgAspLeuGlyProAspGlu 243
DB 726 GCCTCAGGAGCTAGATTAATATAGTACATGATGCGCATTAATCTATCTGCGGGGAG 785
QY 244 GlnSerLeuHisProTrpGluGlnAspLeuLeuLysLysIleCysAspAlaIleTyrTrpLeu 263
DB 786 GAAGCTTTCAGCCGCTGGGAGCAAAACATCTTGCACAAATCTGTAAAGCTTCTATCTC 845
QY 264 ProAlaTrpSerSerThrSerAspTrpValLysLeuGlnSerLeuSerLeuGlnAsp 283
DB 846 CCGGCTGTGGCTCCACCGATGATATGCTCACTGCTTCAATCCATCTCTCCAGGAT 905
QY 284 IleLysSerGluAspTrpSerArgPheValAlaProPheTrpProAlaValIleArgSer 303
DB 906 ATTAAGTGTGCGGATTTGCTCAGAGAACGTAGCTCTTCTGCGCTGGGTTATGCGACT 965
QY 304 AlaPheThrTrpLysGlyLeuSerSerLeuLeuSerSerGlyGlnLysThrIleLysGly 323
DB 966 GCATTACATGAAAGGCGCTTGTCTCTCTCTGCTAGTGTATGAAAGATTAAGAA 1025
QY 324 AlaLeuAlaMetProIleMetIleGlyGlyTyrLysLysAspLeuIleLysPheAlaIle 343
DB 1026 GCATTGTGCAATGCGATTGATGATTGAAGTTACAAAGAGCTGTCATTAACTTTGGTATC 1085
QY 344 IleThrCysArgLysPro 349
DB 1086 ATCACTTGCAGAAAGCA 1103

```

```

RESULT 8
AAK17789
ID AAK17789 standard; DNA; 1790 BP.
XX
AC AAK17789;
XX
XX 21-MAY-1999 (first entry)
XX
XX Arabidopsis gamma-tocopherol methyltransferase encoding gene.
XX
KW Gamma-tocopherol methyltransferase; gamma-TMT; SLR0089; 165H5T7; meat;
transgenic plant; alpha-tocopherol; gamma-tocopherol; forage plant; ds.

```

```

XX OS Arabidopsis sp.
XX
XX FH Key Location/Qualifiers
XX CDS 207..1253
XX FT /tag= a
XX FT /gene= "165H5T7"
XX FT /product= "gamma-TMT"
XX
XX PN W09904622-A1.
XX
XX PD 04-FEB-1999.
XX
XX PF 22-JUL-1998: 98W0-US15137.
XX
XX PR 17-JUL-1998: 98US-0053819.
XX PR 25-JUL-1997: 97US-0053819.
XX PR 26-JAN-1998: 98US-0072497.
XX
XX PA (UYNE-) UNIV NEVADA.
XX
XX PI Dellapenna D, Shintani DK;
XX
XX DR MPI: 1999-142458/12.
XX DR P-PSDB: AAM95017.
XX
XX PT Newly isolated DNA fragment comprising a ^c-tocopherol (vitamin E)
XX PT methyltransferase coding sequence - useful for producing
XX PT ^a-tocopherol, and transgenic plants, seeds and oils with an altered
XX PT tocopherol profile
XX
XX PS Claim 2; Page 36-38; 46pp: English.
XX
XX CC The invention provides DNA sequences encoding gamma-tocopherol
XX CC methyltransferases (gamma-TMT). The gamma-TMT encoding genes SLR0089
XX CC and 165H5T7 are isolated from Synechocystis and Arabidopsis species
XX CC respectively. The DNA fragments are useful for producing transgenic
XX CC plants containing a higher alpha-tocopherol:gamma-tocopherol ratio. They
XX CC are also useful for producing alpha-tocopherol and plants with increased
XX CC gamma-tocopherol, which may be useful in certain industries such as the
XX CC meat industry e.g. for developing forage plants to feed animals. The
XX CC production of transgenic plants (and seeds) with a higher alpha-
XX CC tocopherol:gamma-tocopherol ratio, using the gamma-TMTs, will increase
XX CC the level of alpha-tocopherol in the human diet, in addition to enhancing
XX CC the stability and shelf life of plants and plant products. Increased
XX CC levels of alpha-tocopherol will also increase meat quality and extend
XX CC shelf life of post-processed meat products. Plants with a higher alpha-
XX CC tocopherol:gamma-tocopherol ratio may also produce advantageous
XX CC phenotypes. The present sequence represents the 165H5T7 gene encoding the
XX CC Arabidopsis gamma-TMT.
XX
XX SQ Sequence 1790 BP; 474 A; 404 C; 436 G; 476 T; 0 other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 1,35e-109 Length: 1790
XX Score: 1151.50 Matches: 230
XX Percent Similarity: 75.43% Conserved: 31
XX Best Local Similarity: 66.47% Mismatches: 65
XX Query Match: 62.92% Indels: 21
XX DB: 20 Gaps: 4
XX
XX US-09-857-613a-28 (1-350) x AAK17789 (1-1790)
XX
XX QY 8 ProThrIleSerCysIleHisIleHisThrPheArgSerGlnSerProArgThrPheAla 27
XX DB 259 CCAACTCTCTTTCGGCTCAAGCATCGCTTCTTT-CGGTCCATGCTCTCTCC 317
XX
XX QY 28 ArgIleArgVal---GlyProArgSerTrpAlaProIleArgAlaSerAlaSerSer 46
XX DB 318 TCACCTCTATAGCAGCAACCGCTGGAACCTGCTGGCGCTGCTGCTACATCCACT 377
XX
XX QY 47 GluArgGlyGluIleValLeuGlnGlnLysProLysLysAspAspLysLysLysLeuGln 66
XX DB 111

```

Db 378 GAG-----GGCCTAAGA 389
OY 67 LysGlyIleAlaGluPheThrAspGluSerSerGlyLeuTirGluAsnIleTrrpGlyAsp 86
Db 390 AAGAGAAATGCGGAGCTTCAACAATGAACCTCGGTTGGGAAAGATTTGGAGAT 449
OY 87 HisMetHisHISGlyPheThrAspSerSerThrValSerLeuSerAsp-----His 104
Db 450 CATATGATCATGCGCTTTATATGACCTGATCTCTCTGTTCACTTCTGATTTCTGGTCA 509
OY 105 ArgAlaAlaGlnIleArgMetIleGlnGluSerLeuArgPheAlaSerValSer---Glu 123
Db 510 AAGGAAGCTCAGATCCGTTATGATTAAGAGTCTCCGGTTCCCGGCTTACTGATGA 569
OY 124 GluArgSerLeuTrrpProLysSerIleValAspValGlyCysGlyIleGlyGlySer 143
Db 570 GAGGAGGAGAAAAAGATAAGAAAGAGTGGAGTGGGTTGGGATTTGGAGGAGCTCA 629
OY 144 ArgTyrLeuAlaLysLysPheGlyValThrSerValGlyIleThrLeuSerProValGln 163
Db 630 ACATATCTTGCTCTTAATTTGGAGCTGAATGATGCGCTTACTCTCAGCCTGTTCA 689
OY 164 AlaGlnArgAlaAsnAlaLeuAlaAlaAlaGlnGlyLeuAlaAspLysValSerPheGln 183
Db 690 GCCAAGAGAGCCCAATGATCGCGCGCTCATCTCTCATTAAGCGCTTCTTCA 749
OY 184 ValAlaAspAlaLeuGlnGlnIleProPheSerAspGlyGlnPheAspLeuValTrrpSerMet 203
Db 750 GTTGGCGATGCGTGGATGACCATTCGAAGAGGAAATTCGATGATGCTGCATG 809
OY 204 GluSerGlyGluHisMetProAspLysAlaLysPheValGlyLeuAlaArgValAla 223
Db 810 GAGAGTGGTGAATATGCTGCAAGGCCAAGTTGTAAAGAGTTGATGATGCTGCG 869
OY 224 AlaProGlyAlaIleIleIleValThrTrrpCysHisArgAspLeuGlyProAspGlu 243
Db 870 GCTCCAGAGAGTGAATATATAGTACATGTCGATGAATATCATCTGCGGGGAG 929
OY 244 GluSerLeuHisProTrrpGluGlnAspLeuLysLysIleCysAspAlaTrrpTyrLeu 263
Db 930 GAAGCTTTGAGCCGTTGGGAGCAAAACATCTTGACAAAATCTGTAGACGTTCTATCTC 989
OY 264 ProAlaTrrpCysSerThrSerAspTrrpValLysLeuGlnSerLeuSerLeuGlnAsp 283
Db 990 CCGGCTTGGCTCCACCCGATGATTATGTCACCTTGAATCCCATTTCTCTCAGAGAT 1049
OY 284 IleLysSerGluAspTrrpSerArgPheValAlaProPheTrrpProAlaValIleArgSer 303
Db 1050 ATTAAGTGTGCGGATTTGGTACAGAAAGTACCTCTTGGCGCTGCGGTTATCGGAGCT 1109
OY 304 AlaPheThrTrrpLysGlyLeuSerSerLeuLeuSerSerGlyGlnLysThrIleLysGly 323
Db 1110 GCATTAACATGAGGAGCGCTTGTCTCTCTGCTGCTGATGATAAGATTAAGACA 1169
OY 324 AlaLeuAlaMetProLeuMetIleGlnGlyTyrLysLysAspLeuIleLysPheAlaIle 343
Db 1170 GCATTAGCAATGCGATTTGATGAAGTACAAAGAGTTCATTAGTTGATTC 1229
OY 344 IleThrCysArgLysPro 349
Db 1230 ATCATTGCCAGAACCA 1247

KW protein identification: signal transduction pathway;
KW metabolic pathway: promoter; termination sequence: ss.
XX Arabidopsis thaliana.
OS
XX
PN EP1033405-A2.
PD 06-SEP-2000.
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128334.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132048.
PR 05-MAY-1999; 99US-0132407.
PR 06-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139450.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.


```

QY 119 AlaSerValSerGluGlnArgSerIleTrpProLysSerIleValAlaAspValGlyCysGly 138
DB 559 GCTGGCGTTCCTCATGAGCAAAAGATTGAAACCAACGACATGCTCATGTGGTTGTGG 618
QY 139 IleGlyGlySerSerArgTyrIleuAlaLysLysPheGlyAlaThrSerValGlyIleThr 158
DB 619 ATAGGGGGAAGCTCACGTTACTTGGCCGGAAATTCACAGGCCAAGGTGAATGCCATCAG 678
QY 159 LeuSerProValGlnAlaGlnArgAlaAsnAlaLeuAlaAlaGlnIleuAlaAsp 178
DB 679 CTCAGCCCGTCCAGGTTCAGAGAGCCGTAGACCTTACTGCAACAGCAAGCTTATCTGAC 738
QY 179 LysValSerPheGlnValAlaAspAlaLeuGlnIleProPheSerAspGlyGlnPheAsp 198
DB 739 CTCGTCATATTCAGCTACGATGCAATGCCCTGCAACACCCCTTCAGATGCTTGTGAT 798
QY 199 LeuValTrpSerMetGluSerGlyGlnHisMetProAspLysAlaLysPheValGlyGlu 218
DB 799 CTGCTGTGTCCATGAGAGCGCGGACACATGCCACAGCAAAAAGTTTGTGGCGAG 858
QY 219 LeuAlaArgValAlaAlaProGlnAlaIleIleIleValThrTrpCysHisArgAsp 238
DB 859 CTTCGACGAGTACACACCTCCGCGGCTGCATATTCCTGTGACGTCGTCGACCGCTGAT 918
QY 239 LeuGlyProAspGlnIleSerLeuHisProTrpGlnGlnAspLeuLysIleCys 258
DB 919 CTCAGACCCGCGTGAACCTCTCTCAAGCCTGACGACGACGATCTTTGGCAAGATTGT 978
QY 259 AspAlaTrpTyrLeuProAlaTrpCysSerThrSerAspTyrValLysLeuGlnIleSer 278
DB 979 GACGACTTCTACTGCTGACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1038
QY 279 LeuSerLeuGlnAspIleLysSerGluAspTrpSerArgPheValAlaProPheTrpPro 298
DB 1039 TTAGGCGCTCAGAGATATCAAGAGCGAGGCTGTCGAGTACGTCGACGCTTCTGCGCA 1098
QY 299 AlaValIleArgSerAlaPheThrTrpLysGlyLeuSerSerLeuSerSerGlyGln 318
DB 1099 GCGGTGATGAAACCCGCTTGTCCATGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1158
QY 319 LysThrIleLeuGlnAlaLeuAlaMetProLeuMetIleGlnGlyTyrLysAspLeu 338
DB 1159 ACTACTATGAAAGAGCTTTCCCATGACGCTCATGATCCAGGGCTACACGACGAGGCTC 1218
QY 339 IleLysPheAlaIleIleThrCysArgLys 348
DB 1219 ATTAATTCGCTCCATCCTTCAGAGAG 1248
RESULT 11
AAA29155
ID AAA29155 standard; cDNA; 1011 BP.
XX
XX AAA29155;
AC
XX
XX 12-SEP-2000 (first entry)
DE Wheat gamma tocopherol methyltransferase cDNA.
XX
XX Wheat gamma tocopherol biosynthesis; enzyme; inhibitor; herbicide;
KM Vitamin E; alpha-tocopherol; ss.
XX gamma-tocopherol methyltransferase; ss.
OS
XX
XX Triticum aestivum.
FH
FH Key CDS Location/Qualifiers
FT 3,885
FT /tag a
FT /partial
FT /product- gamma-tocopherol_methyltransferase
FT /transl_except- (pos:381..383, aa:Xaa)
FT /transl_except- (pos:387..389, aa:Xaa)
FT /transl_except- (pos:390..397, aa:Xaa)
FT /transl_except- (pos:393..395, aa:Xaa)

```

```

FT /transl_except- (pos:399..401, aa:Xaa)
FT /note="Xaa is not defined"
PD WO200032757-A2.
XX
XX 08-JUN-2000.
PD 02-DEC-1999; 99WO-US28588.
XX
XX 03-DEC-1998; 98US-0110781.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E. I.
XX
XX Caohon RE, Coughlan SJ, Miao G, Rafalski JA;
XX
XX WPI: 2000-412309/35.
XX
XX P-PSDB: AAY96463.
XX
XX polynucleotide encoding a vitamin E biosynthetic enzyme polypeptide
XX such as gamma tocopherol methyltransferase, useful for synthesizing
XX vitamin E or as a target for design and discovery of herbicides
XX
XX Claim 3: Page 53; 82pp; English.
XX
XX AAA29151-69 encode vitamin E (alpha-tocopherol) biosynthetic enzymes or
XX their fragments. The cDNA clones were identified by BLAST searches based
XX on similarity to Synchocystis sp. and Arabidopsis thaliana vitamin E
XX biosynthetic enzymes. The enzymes are useful for synthesizing
XX plastoquinones or tocopherols (especially vitamin E). Vitamin E is
XX required for plant growth, therefore the enzymes may be used for the
XX discovery of new herbicides. The enzymes can be used in methods to
XX evaluate potential inhibitors, which may have use as herbicides.
XX Additionally, transgenic expression of, e.g. gamma-tocopherol
XX methyltransferase affords the ability to manipulate tocopherol levels as
XX desired for a particular application.
XX
XX Sequence 1011 BP; 255 A; 269 C; 268 G; 216 T; 3 other:
XX
XX Alignment Scores:
XX Pred. No.: 1,24e-89 Length: 1011
XX Score: 956.00 Matches: 194
XX Percent Similarity: 79.21% Conservative: 27
XX Best Local Similarity: 69.53% Mismatches: 55
XX Query Match: 52,248 Indels: 4
XX DB: 21 Gaps: 0
XX
XX US-09-857-613A-28 (1-350) x AAA29155 (1-1011)
XX
XX QY 75 GluSerSerGlyLeuTrpGluAsnIleTrpGlyAspHisMetHisGlyPheTyrAsp 94
XX DB 3 GACTGCTCCGCGCTGTGGAGACATCTGGGGGAGACATGACCAACGCGCTTCTACGAC 62
XX
XX QY 95 SerAspSerThrValSerLeuSerAspHisArgAlaAlaGlnIleArgMetIleGlnIle 114
XX DB 63 TCCGGCGAGCGCCGCTTCATGTCGACACCGCCGCGCCACATGCCATGTGAGAG 122
XX
XX QY 115 SerLeuArgPheAlaSerValSerGluGlnArgSerIleTrpProLysSerIleValAsp 134
XX DB 123 GCCCTGCGCTTCCGCGCGCTCCCGACGATCCCAACAAACCAACCAACGATGTGAT 182
XX
XX QY 135 ValGlyGlyGlyIleGlyGlySerSerArgTyrLeuAlaLysLysPheGlyAlaThrSer 154
XX DB 183 GTTGGATGCGGAATCGGTGACTCATAGATACCTGGGCAACAAATATGAGCAATG 242
XX
XX QY 154 rValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAsnAlaLeuAlaAlaGln 174
XX DB 243 CTCTGGATCATGTCAC-CCAGTGCAGCTGAGAGAGAAATGCCCTCCGCGGAGCGCA 301
XX
XX QY 174 nGlyLeuAlaAspLysValSerPheGlnValAlaAspAlaLeuGlnIleProPheSerAs 194
XX DB 302 AGGGGTGTCCGGAACAAGTTCTTCCATATTCGATCTCGGAGCAACCAATTTCTCTCG 361
XX
XX QY 194 pGlyGlnPheAspLeuVal-TrpSerMetGluSerGlyGlnHisMetProAspLysAla 214

```

```
Db 362 ATGGCCATTGATCTTTGTCGGGNCANTGGAGANTGGTGACACATGCGCAACAAACAGA 421
Qy 214 ysphevalglvleuvalaargvalaalaaproglialaallellellelvalthr 234
Db 422 AGTTGTGAAGGAGCTGGCAGCGCTCGCAGCTCCAGAGAACATCATCATCGTGACCT 481
Qy 234 rrcysahisargaspheuglyproaspgluglnserleuhisprotrpoglinspleu 254
Db 482 GGTGCATAGAGAACCTCGCCCATCGAGAGACTGACGAACTCAGACAGCTGAATCTT 541
Qy 254 eulyslylleucysasplalatyrryleuproalaatrrpcysserthseraspyrrval 274
Db 542 TGAAGAAAGATTGTGATGATATATATACCTCCGATTTGGCTCTCCCTCGGATTTATGCA 601
Qy 274 yslleuclnserleuSerleuGlnaspilleySerleuAsprrpserarphava 294
Db 602 AGATTGCGGAGCATTTGCTCTTTAGAGATATCAAAAGCCGACAGTGTCAAAAGCTGG 661
Qy 294 lairpohetrrproalavalileatrgseralaphethrrtrpysglyleuSerleu 314
Db 662 CCCCTTCTGGCTCTGTCATCATCATGACAGCAGTCAATGAAAGCCCTCACCTCTCTAC 721
Qy 314 euserSerleuglnlystrhrilleysglyalaleuAlameprroeuMetileugly 334
Db 722 TTAGAGATGTGATGAGAGACATTAAGGAGCAGCTGTGATGCTCTCATGATCAAGCT 781
Qy 334 yrllyslyaspleuileysphe-AlallellethrCysarglysProglu 350
Db 782 ACAAGAAAGGCTCATTTAGTTTCAAGCATCATCACTGCGCAAAACCCCAA 832

RESULT 12
AAA29151
ID AAA29151 standard: cDNA: 792 BP.
AC AAA29151:
XX 12-SEP-2000 (first entry)
DE Corn gamma tocopherol methyltransferase partial coding sequence.
XX
XX Vitamin E: alpha-tocopherol; biosynthesis; enzyme; inhibitor; herbicide;
KW gamma-tocopherol methyltransferase; ss.
XX
XX Zea mays.
OS
XX
XX Key Location/Qualifiers
FH 2..574
FT CDS /tag= a
FT /product= gamma_tocopherol_methyltransferase
FT /partial
XX
XX MO200032757-A2.
PD 08-JUN-2000.
XX
XX PF 02-DEC-1999; 99WO-US28588.
XX
XX PR 03-DEC-1998; 98US-0110781.
XX
XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX PI Cahoon RE, Coughlan SJ, Miao G, Rafalski JA;
XX
XX DR WPI: 2000-412309/35.
XX
XX P-PSDB: AAY96459.
XX
XX Poly nucleotide encoding a vitamin E biosynthetic enzyme polypeptide
XX such as gamma tocopherol methyltransferase, useful for synthesizing
XX vitamin E or as a target for design and discovery of herbicides
XX
XX Claim 3; Page 49; 82pp; English.
XX
```

```
CC AAA29151-69 encode vitamin E (alpha-tocopherol) biosynthetic enzymes or
CC their fragments. The cDNA clones were identified by BLAST searches based
CC on similarity to Synchocystis sp. and Arabidopsis thaliana vitamin E
CC biosynthetic enzymes. The enzymes are useful for synthesizing
CC plastoquinones or tocopherols (especially vitamin E). Vitamin E is
CC required for plant growth, therefore the enzymes may be used for the
CC discovery of new herbicides. The enzymes can be used in methods to
CC evaluate potential inhibitors, which may have use as herbicides.
CC Additionally, transgenic expression of, e.g. gamma-tocopherol
CC methyltransferase affords the ability to manipulate tocopherol levels as
CC desired for a particular application.
XX
SQ Sequence 792 BP; 201 A; 204 C; 215 G; 166 T; 6 other:

Alignment Scores:
Pred. No.: 2.77e-71 Length: 792
Score: 778.00 Matches: 145
Percent Similarity: 85.86% Conservative: 19
Best Local Similarity: 75.92% Mismatches: 27
Query Match: 42.51% Indels: 0
Db: 21 Gaps: 0

US-09-857-613a-28 (1-350) x AAA29151 (1-792)
Qy 159 leuserprovalglinalglinalrvalaalsnaalaleuAlaAlaaglnclyleuAlasp 178
Db 2 TTGAGCCCTGTTCAAGCCGAGAGAGAAATGCTTCCTCGCCAGCCGAGGGTTGCGCAT 61
Qy 179 lysvalserphelglvalaAlaasplaleuGlnInpropheseraspolyglinsp 198
Db 62 CAGGTTACTCTGCAGAGTTGCTGATGCTGAGCAACGCTTCCTGAGCGGAGTTGCAT 121
Qy 199 leuvaltrpsermetgluserglylunhismetproaspilyalalyshelvalglu 218
Db 122 CTGCTGTGCTCATGAGAGAGTGGCGAGCAGCATGCCGAGAAAGATTGTGTAGAG 181
Qy 219 leuAlaargvalaAlaAlaaproglyalallellellellelvalThrTrpCysHisargasp 238
Db 182 CTAGCAGCGGTGGCGGCTCTCTGAGGAGCAATATATCATGTGATGCTCATGAGAAC 241
Qy 239 leuglyproaspgluglnserleuhisprotrpoglinspleuileuylslylecy 258
Db 242 CTGATCCATCCGAAACCTCGCTAAAGCCGATGAACCTGAGCTCTGAGAGATATGC 301
Qy 259 aspalaatyrryleuproalatrppysserthrseraspyrrvalysleuclnser 278
Db 302 GACGCTACTACCTCCGCGAGCTGCTCACCTTCAGACTATGTGAACATTCAGACTCA 361
Qy 279 leuserleuGlnaspilleySerleuAsprrpserarphavaAlaIrrpohetrrpro 298
Db 362 CTGCTCTCGAGGATATATCAAGACAGCTGACTGTCGAGACAGTGGCCCGTTTG6CCC 421
Qy 299 AlavalilearSerleuThrtrpysglyleuSerleuSerleuSerleuSerleugln 318
Db 422 GCCGTATATAATCAAGCGGTATACATGAGAGGCTTACCTCTCTGTCGACAGCCGATG 481
Qy 319 lysthrilleysglyalaleuAlameprroeuMetileuglytyrrylslyaspleu 338
Db 482 AAGACGATCAGAGCGCGATGATGATGATATATGATCAGGCGTACAAAGAGGGGCTC 541
Qy 339 lleysphealallellethrCysarglyspro 349
Db 542 ATCAAAATTCACCATCATCATCTGCGAAGCCT 574

RESULT 13
AAK17788
ID AAK17788 standard: DNA; 954 BP.
XX
XX AAK17788:
XX
XX 21-MAY-1999 (first entry)
XX
XX Synchocystis gamma-tocopherol methyltransferase encoding gene.
XX
```


XX DE19931834-A1.
 XX 11-JAN-2001.
 XX 09-JUL-1999; 99DE-1031834.
 XX 09-JUL-1999; 99DE-1031834.
 XX 09-JUL-1999; 99DE-1031834.
 XX (SUNG-) SUNGENE GMBH & CO KGAA.
 XX Badur R, Herbers K, Kunze I, Geiger M, Mock H:
 XX WPI; 2001-148279/16.
 XX P-PSDB; AAB46833.
 XX New DNA encoding 2-methyl-6-phytylhydroquinone methyltransferase,
 XX useful for producing transgenic plants with increased content of
 XX tocopherols and tocotrienols.
 XX Claim 1; Page 20-22; 28pp; German.
 XX This invention describes a novel DNA (I) which encodes a
 XX 2-methyl-6-phytylhydroquinone methyltransferase, MPMT (a) from
 XX *Synechocystis* and 1s: (i) sequence (a) or (b), of 957 and 930 base pairs
 XX (bp) respectively, given in the specification; (ii) sequence that
 XX hybridizes to (i), or (iii) is homologous to all or part of (i). The
 XX invention also describes (1) use of DNA sequences encoding an (A) for
 XX producing plants or photosynthetic organisms with increased
 XX concentrations of tocopherols (ii) and tocotrienols (iii); (2) plants
 XX with increased contents of (ii) and (iii) that include an expression
 XX cassette containing (i); and (3) test kit, based on expression of (a),
 XX (b) or sequences that hybridize to them, for identifying inhibitors of
 XX (A). (i), or generally any DNA that encodes (A), are used to produce
 XX transgenic plants or other photosynthetic organisms that have increased
 XX contents of tocopherols and tocotrienols (vitamin E). (A) catalyses
 XX biosynthesis of 2,3-dimethyl-6-(phytyl or geranylgeranyl) hydroquinone,
 XX intermediates for tocopherols and tocotrienols. (a), (b) or sequences
 XX that hybridize with them can also be used to identify inhibitors of (A).
 XX
 XX Sequence 930 BP; 205 A; 230 C; 262 G; 233 T; 0 other:
 XX
 XX Alignment Scores:
 XX Pred. No.: 4.51e-31 Length: 930
 XX Score: 392.00 Matches: 92
 XX Percent Similarity: 51.76% Conservative: 40
 XX Best Local Similarity: 36.08% Mismatches: 87
 XX Query Match: 21.42% Indels: 36
 XX DB: 22 Gaps: 10
 XX
 XX US-09-857-613A-28 (1-350) x AAF26173 (1-930)
 OY 69 IIEAIAIGLPHETRYASPGLU-----SERSERGLYLEUTRPGIUNSNLIERPGIYASP 86
 DB 64 GTGGCCACCGCTACGACCAATGACAGACAGACGATTTGGCAATATTACTGGGGCGAC 123
 OY 87 HSMETHLSHISGLYPHETRYASPERASPERTHRVASERLEUSERASPHISARGALA 106
 DB 124 CATATCCACCTCGGCATTTAT---GGCGATCCGCCAGTGGCCAGAT----- 168
 OY 107 AIAGLINLEARGMETLIEGLINSEURLEARGPHEALIASERVALSERGLUARGSER 126
 DB 169 -----TTCAATCCATCCAAATGATTTGTC-----CATGCCATGCC 207
 OY 127 LYSTTP-----PRLYSERLIEVALASPYALIGLYSGLY 138
 DB 208 CAGTGGGGCGGATTACACTCTCCCGCCGACACAGGTATGTGATGGGTGGCGC 267
 OY 139 IIEGLYGLSERSEARGTYLEUALATYLSYPHEGLIALATHRSERVALIGLYLIETHR 158
 DB 268 ATTTGGGGGAGAGTGCATTTCTCGCCAAAGATATGCTTTTAACTTACCGGCATCACC 327
 OY 159 LEUSERPROVALGALNALAGINALARGALASNALALALALAGLINCGLYLEUALIASP 178

DB 328 ATTAGTCCCAACAGGTGAACGGCGACGCGATTAACCTCCCGCATGTGACGCC--- 384
 OY 179 LYSVALSERPHEGLINVALALASPALALEUNGINGINPROPHESERASPGIYNPHEASP 198
 DB 385 -----AAGTTTGGCGGTGACGATCTATGCTTTGCTTTTCCAGAGTAGTTTCAC 438
 OY 199 LEUVALTRPSERMETGLUSERGLYUHSMETPROASPLYALALYSPHEVALIGLYLU 218
 DB 439 GTAGTTTGGTCGCTGGAAGCAGGCCCCACATGCTGACAAAGCTGTGTTGCCAAGGAA 498
 OY 219 LEUVALARGVALALALAPROGLYALALLELLELLELLEVALTHRTPCSHISARGASP 238
 DB 499 TTACTGCGGGCTGGAACACGAGGGGCTGCTGTGGTGGCGGATTCATCAACAGGAC 558
 OY 239 LEUGLYPROASPGIUNL---SERLEUHSPTOTRPGIUNASPLEUENLYSLISLE 257
 DB 559 -----GATCGCCAACTGCCCCCTCAACTCTGGGAAAAACAGTAGATGGACAACTG 609
 OY 258 CYSASPALATRYTYRLEUPROALATRPCYSERTHRSERASPTRYVALLYSLEUENGLN 277
 DB 610 TTGATCAATGCTCCACCCCTCTTCCACCATTTGAAGCTTTTGGCGAAATTTGGAA 669
 OY 278 SERLEUSERLEUGLNASP-----ILEYSSERGLUASPTRPSERARGPHEVALALAPRO 295
 DB 670 GCCACGGCTTGTGGAGGGCCAGGTGACTACTGCTGATTGACTGTACCGACCTCCCC 729
 OY 296 PHEPTRPROALAVALLIARGSERALAPHERNTRPYLSGLYLEU 310
 DB 730 GCTGTGTTGGATACCATV-----TTGCACGGCATTT 759
 XX
 XX RESULT 15
 XX AAZ61599
 XX ID AAZ61599 standard; DNA; 957 BP.
 XX XX
 XX AC AAZ61599;
 XX XX
 XX DT 19-JUN-2000 (first entry)
 XX XX
 XX DE DNA encoding a methyltransferase 1 protein.
 XX XX
 XX KW 2-methyl-6-phytylplastoquinol methyltransferase; plastoquinol;
 XX KW 2-methyl-6-solanyplastoquinol-9 methyltransferase; methyltransferase 1;
 XX KW *Synechocystis* strain PCC6803; tocopherol; transgenic plant;
 XX OS delta-tocopherol; gamma-tocopherol; nutritive value; ss.
 XX XX
 XX Synchocystis sp.
 XX XX
 XX FH Key Location/Qualifiers
 XX FT CDS 1..957
 XX FT /*tag= a
 XX FT /product= "2-methyl-6-phytylplastoquinol/2-methyl-
 XX 6-solanyplastoquinol-9 methyltransferase"
 XX PN WO200010380-A1.
 XX XX
 XX PD 02-MAR-2000.
 XX XX
 XX PF 25-AUG-1999; 99WO-US19483.
 XX XX
 XX PR 25-AUG-1998; 98US-0097863.
 XX XX
 XX PA (UYNE-) UNIV NEVADA.
 XX XX
 XX PI Dellapenna D, Shintani DK;
 XX XX
 XX DR WPI; 2000-224497/19.
 XX XX
 XX DR P-PSDB; AAY69405.
 XX XX
 XX Genetic construct comprising a
 XX 2-methyl-6-phytylplastoquinol/2-methyl-6-solanyplastoquinol-9
 XX methyltransferase coding sequence, useful for altering the ratio of
 XX delta-tocopherol : gamma-tocopherol in plants -

XX Claim 2; Page 30-32; 34pp; English.
PS
XX The present sequence encodes a 2-methyl-6-phytylplastoquinol/2-methyl-
CC 6-solanyplastoquinol-9 methyltransferase (methyltransferase 1) protein.
CC The enzyme isolated from *Synechocystis* strain PCC6803. The enzyme is
CC a fundamental enzyme in the production of tocopherols and plastoquinols
CC in higher plants. The enzyme is believed to be involved in regulating
CC the relative amounts of the various tocopherols in photosynthetic
CC organisms. Genetic constructs comprising the methyltransferase 1 coding
CC region under the control of a plant promoter are used to produce
CC transgenic plants, to alter the ratio of delta-tocopherol:gamma-
CC tocopherol in plants. This increases the nutritive value of the plants
CC and products produced from them for humans and animals.
XX

SO Sequence 957 BP; 203 A; 244 C; 268 G; 242 T; 0 other;

Alignment Scores:

Pred. No.:	4,7e-31	Length:	957
Score:	392.00	Matches:	92
Percent Similarity:	51.76%	Conservative:	40
Best Local Similarity:	36.08%	Mismatches:	87
Query Match:	21.42%	Indels:	36
DB:	21	Gaps:	10

US-09-857-613a-28 (1-350) x AA261599 (1-957)

OY 69 IleAlaGluPheTyrAspGlu-----SerSerglyLeuTrpGluAsnIleTrpGlyAsp 86
DB 106 GTGGCCAAAGCGCTTACGACCAATGACAGACAGCGCATTTGGAAATATTACTGGGGCCAC 165
OY 87 HlsmethlshlglyPheTyrAspSerSerpThrValSerLeuSerAspHisArgAla 106
DB 166 CATATCCACCTCGGCATTAT---GGCATCCGCGCACATGCCAAGCAT----- 210
OY 107 AlaGlnIleArgMetIleGlnGluSerLeuArgPheAlaSerValSerGluGluArgSer 126
DB 211 -----TTCATCCAAATCGAAATGATTTGTC-----CATGCCATGGCC 249
OY 127 LysTrp-----ProlYsSerIleValAspValGlyCysGly 138
DB 250 CAGTGGGGCGGATTAGATACACTTCCCGCGCACACGCTATTGGATGGGTGGCCGC 309
OY 139 IleGlyGlySerSerpArgTyrLeuAlaLysLysPheGlyAlaThrSerValGlyIleThr 158
DB 310 ATTGGCGGTAGACGTCGATTCTCGCAAGATATGTTTACGTTACCGCATCAAC 369
OY 159 LeuSerProValGlnAlaGlnArgAlaAsnAlaLeuAlaAlaGlnGlyLeuAlaAsp 178
DB 370 ATTAGTCCCAACAGTGAACAGGGCGACGGAATTACTCTCCCGATGTGACGGCC--- 426
OY 179 LysValSerPheGlnValAlaAspAlaLeuGlnInProPheSerAspGlyGlnPheAsp 198
DB 427 -----AAGTTTGGCGTGAACATGCTAAGCTTTGCTTCTTCTGACGAGTATTCCGAC 480
OY 199 LeuValTrpSerMetGluSerGlyGluHisMetProAspLysAlaLysPheValGlyGlu 218
DB 481 GTAGTTTGTGCTGAGAGCAGGGCCCAATGCTGACAAAGCTGTGTTGCCAAGGAA 540
OY 219 LeuAlaArgValAlaAlaProGlyAlaIleIleIleIleValThrTrpCysHisArgAsp 238
DB 541 TTACTGCGGTCGTGAACACAGGGGCGCATTTGCTGTGGCGATTGGAATCAACGGGAC 600
OY 239 LeuGlyProAspGluGln---SerLeuHisProTrpGluGlnAspLeuLeuLysIle 257
DB 601 -----GATGCCCAAGTCCCTCAACTCTGGGAAAAACAGATGATGCGACACATG 651
OY 258 CysAspAlaTyrTrpLeuProAlaTrpCysSerThrSerAspTyrValLysLeuGln 277
DB 652 TTGGATCAATGCTCCACCTCTTTCGACACATTTGAAGTTTTCGCAAAATTTGGA 711
OY 278 SerLeuSerLeuGlnAsp-----IleLysSerGluAspTrpSerArgPheValAlaPro 295

DB 712 GCCACGGCTTTGGTGAGAGGCCAGGTGACTACTGCTGATTGGACTGTACCGACCTTCCCC 771

OY 296 PheTrpProAlaValIleArgSerAlaPheThrTrpLysGlyLeu 310

DB 772 GCTTGGTTGGATACCATT-----TGCAGGGCAT 801

Search completed: March 19, 2003, 07:20:17
Job time : 234 secs

13	131	7.2	8051	2	US-08-576-626A-2	Sequence 2, Appl
14	127	6.9	4403765	4	US-09-103-840A-2	Sequence 2, Appl
15	120	6.6	3719	1	US-08-920-812-10	Sequence 10, Appl
16	120	6.6	3719	1	US-08-920-812-10	Sequence 10, Appl
17	120	6.6	3719	1	US-08-921-177-10	Sequence 10, Appl
18	120	6.6	3719	1	US-08-362-577C-10	Sequence 10, Appl
19	120	6.6	3719	2	US-08-920-828-10	Sequence 10, Appl
20	117.5	6.4	4411529	4	US-09-103-840A-1	Sequence 1, Appl
21	114.5	6.3	4451	3	US-09-217-609A-21	Sequence 21, Appl
22	114.5	6.3	2451	1	US-08-873-235B-21	Sequence 21, Appl
23	114.5	6.3	4403765	4	US-09-103-840A-2	Sequence 2, Appl
24	113.5	6.2	741	3	US-09-217-609A-29	Sequence 29, Appl
25	113.5	6.2	741	4	US-08-873-235B-29	Sequence 29, Appl
26	109	6.0	9377	4	US-09-221-017B-1002	Sequence 1002, Ap
27	108.5	5.9	2077	4	US-09-221-017B-821	Sequence 821, App
28	108.5	5.9	4463	2	US-08-760-489-1	Sequence 1, Appl
29	108.5	5.9	4463	2	US-08-760-489-3	Sequence 3, Appl
30	108.5	5.9	4463	4	US-09-185-373-1	Sequence 1, Appl
31	108.5	5.9	4463	4	US-09-185-373-3	Sequence 3, Appl
32	107.5	5.9	8257	4	US-09-484-970B-65	Sequence 65, Appl
33	104	5.7	756	2	US-08-401-068-11	Sequence 11, Appl
34	104	5.7	756	2	US-08-846-338-11	Sequence 11, Appl
35	104	5.7	5872	3	US-08-411-768B-1	Sequence 1, Appl
36	104	5.7	5872	3	US-08-411-768B-6	Sequence 6, Appl
37	103	5.6	4911	4	US-09-718-692-1	Sequence 1, Appl
38	103	5.6	4911	4	US-09-718-852-1	Sequence 1, Appl
39	103	5.6	4911	4	US-09-718-851-1	Sequence 1, Appl
40	102.5	5.6	5100	1	US-08-457-245-1	Sequence 1, Appl
41	102.5	5.6	4411529	4	US-09-103-840A-1	Sequence 1, Appl
42	101.5	5.5	5822	3	US-08-899-595-4	Sequence 4, Appl
43	101.5	5.5	5822	3	US-08-899-595-5	Sequence 5, Appl
44	101	5.5	1406	4	US-08-936-165A-119	Sequence 119, App
45	99.5	5.4	777	4	US-09-134-001C-407	Sequence 407, App

```
RESULT 1
US-09-382-906A-1
; Sequence 1, Application US/09382906A
```

```

1 Patent NO. 64494/75
2
3 GENERAL INFORMATION:
4
5 APPLICANT: Dellapenna, Dean
6
7 APPLICANT: Shintani, David
8
9 TITLE OF INVENTION: Manipulation of Tocopherol Levels in Transgenic Plants
10
11 FILE REFERENCE: 920905.90032
12
13 CURRENT APPLICATION NUMBER: US/09/382,906A
14
15 CURRENT FILING DATE: 1999-08-25
16
17 PRIOR APPLICATION NUMBER: 60/097,863
18
19 PRIOR FILING DATE: 1998-08-25
20
21
22 NUMBER OF SEQ. ID NOS.: 4
23
24 SOFTWARE: PatentIn Ver. 2.1

```

```

? SEQ ID NO 1
? LENGTH: 957
? TYPE: DNA
? ORGANISM: Synechocystis PCC6803
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)..(954)
?
? ..282,806,1

```

Alignment Scores:	
Pred. No.:	6,31e-17
Score:	392.00
Percent Similarity:	51.76%
Best Local Similarity:	36.08%
Query Match:	21.42%
DB:	4
Length:	955
Matches:	92
Conservative:	40
Mismatches:	87
Indels:	16
Gaps:	10

US-09-857-613A-28 (1-350) x US-09-382-906A-1 (1-957)

```

69 ILEAGLPHETyrAspGlu-----SerSerGlyLeuTrpGluAsnIleTrpGlyAsp 80
:::|:| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |

```

Db	106	GTGGCAACGCCCTACAGCAACAATGACAGAGACGCGATTTTGGAAATTTACTGGGGCAG	165
Qy	87	HismethIshISglYpHeThYAspSerAspSerThYAlSerLeuSerAspHisAla	106
Db	166	CATATCCACCTCGGCCATTAT--GGCGATCCGCCAGTGGCGAAGAT-----	210
Qy	107	AlaGlnIleArGmetIleGlnIleSerLeuArpHeAlaSerValSerGlnIleArGSer	126
Db	211	-----TTTATCCAAATCGAAAATTTGATTTC-----CATGCCATGGCC	249
Qy	127	LySTrp-----ProlYSerIleValAspValGlyCySgIy	138
Db	250	CAGTGGGGCGGATAGATACACTTCCCCCGGCACACGCGATTGATGTGGTGGCG	309
Qy	139	IleGlyGlySerSerArGTYrLeuAlaLysLysPheGlyAlaThrSerValGlyIleThr	158
Db	310	ATTGGCGGTACACTGCCATTCTCGCAAAAGATTATGTTTAACTTACCGCATCACC	369
Qy	159	LeuSerProValGlnAlaGlnArgAlaAsnAlaLeuAlaAlaAlaGlnGlyLeuAlaAsp	178
Db	370	ATTATGTCCTCCCAACAGGTGAACGGGGCAGCGGAATTAATCTCTCCGATGTGACGCC--	426
Qy	179	LysValSerPheGlnIleAlaAspAlaLeuGlnIleProPheSerAspGlyGlnPheAsp	198
Db	427	-----AAGTTGGCGGTGAGAGATGCGTATGGCTTTGCTTTCTGTACGAGTAGTTTGAC	480
Qy	199	LeuValTrpSerMetGluSerGlyLHisMetProAspLysAlaLysPheValGlyL	218
Db	481	GTACTTGGTGGTGGTGAAGCAGGCGCCACCATTCGCTCGCAAAAGCTGTTGGCCAAAGA	540
Qy	219	LeuAlaArgValAlaAlaProGlyAlaIleIleIleIleValIleTrpTrpCysHisArgAsp	238
Db	541	TTTATCGCGGTGCGAAACCGAGGGGCAATCTGTGTGGCGGATTTGAATCAACGGGAC	600
Qy	239	LeuGlyProAspGlnGln--SerLeuHisProTrpGlnGlnAspLeuLeuLysIle	257
Db	601	-----GATGCCCAAGTGGCCCTCACTTCTGGGAAAAACAGTGAATCCGACAACTG	651
Qy	258	CysAspAlaTrpTYrLeuProAlaTrpCysSerThrSerAspTYrValLysLeuGln	277
Db	652	TTGGATCAATGCTCCACACCTGCGCTTTGCCAGCATTTGAAGTTTTCGGGAAATTGGAA	711
Qy	278	SerLeuSerLeuGlnAsp-----IleLysSerGluAspTrpSerArGpHeValAlaPro	295
Db	712	GCCACGGGTTTGGTGAAGGGCGACGACTCACTCTGATTTGGACTGTACCGACCCCTCCC	771
Qy	296	PheTrpProAlaValIleArgSerAlaPheTrpTrpGlyLeu	310
Db	772	GCTTGGTTGGATACCAAT-----TGGCAGGGCAATT	801

```

; NAME/KEY: misc_RNA
; LOCATION: (1)-(2122)
; OTHER INFORMATION: product = 2.1 kb region
US-09-029-603-1

Alignment Scores:
Pred. No.: 7,81e-21 Length: 2122
Score: 261.00 Matches: 80
Percent Similarity: 46.59% Conservative: 50
Best local Similarity: 28.67% Mismatches: 113
Query Match: 14,268 Indels: 36
DB: 4 Gaps: 10

```


FILING DATE: 09-MAR-1998
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Stuart, Donald R
 REGISTRATION NUMBER: 28,479
 REFERENCE/DOCKET NUMBER: 50,608
 TELEPHONE: (317)337-4816
 TELEFAX: (317)337-4847
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 80161 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-09-036-987A-1

Alignment Scores:
 Pred. No.: 1.73e-13 Length: 80161
 Score: 220.50 Matches: 78
 Percent Similarity: 43.99% Conservative: 50
 Best Local Similarity: 26.80% Mismatches: 125
 Query Match: 12.05% Indels: 39
 DB: 3 Gaps: 9

US-09-857-613a-28 (1-350) x US-09-036-987A-1 (1-80161)

QY 69 ILEAGLUPHETRYRASPGLUSERSERGLYLEUTRPGIUSNILETRPGIYASP----- 86
 DB 20201 GTTGGCGGAGATGATGACCTGTCACCGCGCTTGCAGACTCGGTGCGGGCGCCCTGC 20260
 QY 87 HSMETHISHISGLYPHETRYRASPSPERTHRVALSERLEUSERSPHISARGALA 106
 DB 20261 GCCATCCACCGCTGCTGAGACGACGCGCGCTTCTGTCAGCAGCC-CCGCGA 20319
 QY 107 ALAGNILEARGMETILEGLIUSERLEUARGPHEALASERVALSERGLUGIUSR 126
 DB 20320 CCGGCTACCCGACCTTGTGCGGACG-----GACCGTGT 20335
 QY 127 LYSTPRPOLYSER-ILEVALASPVALGICYSGIYILEGLYISERSEARFTRYLE 146
 DB 20356 CGATGGCGGCGGTGACGTCGATGCGGTGCGGTACCGACACACGCGCTCGCGT 20415
 QY 146 WALALYSLYSPHEGLYALATHRSERVALGILYTLERHLEUSERPROVALGIALAGLNR 166
 DB 20416 CGCGCGACACACCGCATCCAGATCACCAGCGCATCCCTGACAGTGGCCAT 20475
 QY 166 GALASNALEUALAALAGLNGLYLEUALASPLYVALSERPHEGLINVALALAS 186
 DB 20476 CGCCCTGATTCGCGACCGGACGCGGACTAGCCACCGGCTCTCTCGTGGTCTGA 20535
 QY 186 PALALEUNGINGINPROPHESERASPLYGLINPHEASPLEUVALTRPSEMETGLUSER 206
 DB 20536 TGGCATTCCTGCGCGTCCGCGACATGCTTCACGCGCGCTGCGCATCGATCGCT 20595
 QY 206 YGLUHISMETPROASPLYALALYSPHEVALGILYLEUALAARGVALALALAPROGL 226
 DB 20596 GTTGAAGATGTCGACACCGGACGCGTCCGGAATCTTCAGPACTCAACCCGG 20655
 QY 226 YALALELLELLELLELLELLELLELLELLELLELLELLELLELLELLELLELLEL 245
 DB 20656 TGGCATTCCTGCGCGTCCGCGACGCGTCAACAGGAGAGCGGCGGATCCGGTCTC 20715
 QY 245 RLEUHSIPROTPGLUGIUSPLEUVALYSLYILECYASAPALATYTRYLEUPROAL 265
 DB 20716 CGGCGACAGGTGCGCGCGCTT-----CGGATTCCTCGGCT----- 20755
 QY 265 ATRPCYSERTHRISERASPLYRVALYLSLEULEUNGINSERLEUSERGLIASPILLEY 285
 DB 20756 -----GAGCACTTCTGGAATGCTGCGGACGCGGCGGTTCGA 20793
 QY 285 SSERGLIASPTTP-----SERARGPHEVALALAPROPHETPROAL 299

DB 20794 GATCCTCATTTGGAGGACGTCGTCGACGACCGGTACTTCATGCGCGAGTTCGCCGA 20853
 QY 299 AVALLEARGSERALAPERTHRTPLYSGIYLEUSERSERLEUSERSEGLYELNLY 319
 DB 20854 AGAGCTC-----GCTGGCACACACCGGATCCGACAGATGACGG-----CC 20898
 QY 319 STHRILEYSGIYALALEUALAMETPROLEUMETILEGLIYTRYLYLSASPLEUL 339
 DB 20899 GGCTGTGCGCGGCTGGCGCGCGCGCTGCGATTTATGAAATATGCCCCACGATG-- 20956
 QY 339 ELYSPHEALALELLELLELLELLELLELLELLELLELLELLELLELLELLELLEL 349
 DB 20957 -GGCTATGCGATTCGACGCGCGGACGCGC 20986

RESULT 5
 US-09-370-700-1
 : Sequence 1, Application US/09370700
 : Patent No. 6274350
 : GENERAL INFORMATION:
 : APPLICANT: Balitz, Richard H
 : APPLICANT: Broughton, Mary C
 : APPLICANT: Crawford, Kathryn P
 : APPLICANT: Madduri, Krishnamurthy
 : APPLICANT: Treadway, Patil J
 : APPLICANT: Turner, Jan R
 : TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
 : FILE REFERENCE: 50489 DIV1
 : CURRENT FILING DATE: US/09/370,700
 : EARLIER FILING DATE: 1999-08-09
 : EARLIER APPLICATION NUMBER: US 09/36987
 : NUMBER OF SEQ ID NOS: 39
 : SOFTWARE: Patent In Ver. 2.0
 : SEQ ID NO 1
 : LENGTH: 80161
 : TYPE: DNA
 : ORGANISM: Saccharopolyspora spinosa
 US-09-370-700-1

Alignment Scores:
 Pred. No.: 1.73e-13 Length: 80161
 Score: 220.50 Matches: 78
 Percent Similarity: 43.99% Conservative: 50
 Best Local Similarity: 26.80% Mismatches: 125
 Query Match: 12.05% Indels: 39
 DB: 4 Gaps: 9

US-09-857-613a-28 (1-350) x US-09-370-700-1 (1-80161)

QY 69 ILEAGLUPHETRYRASPGLUSERSERGLYLEUTRPGIUSNILETRPGIYASP----- 86
 DB 20201 GTTGGCGGAGATGATGACCTGTCACCGCGCTTGCAGACTCGGTGCGGGCGCCCTGC 20260
 QY 87 HSMETHISHISGLYPHETRYRASPSPERTHRVALSERLEUSERSPHISARGALA 106
 DB 20261 GCCATCCACCGCTGCTGAGACGACGCGCGCTTCTGTCAGCAGCC-CCGCGA 20319
 QY 107 ALAGNILEARGMETILEGLIUSERLEUARGPHEALASERVALSERGLUGIUSR 126
 DB 20320 CCGGCTACCCGACCTTGTGCGGACG-----GACCGTGT 20335
 QY 127 LYSTPRPOLYSER-ILEVALASPVALGICYSGIYILEGLYISERSEARFTRYLE 146
 DB 20356 CGATGGCGGCGGTGACGTCGATGCGGTGCGGTACCGACACACGCGCTCGCGT 20415
 QY 146 WALALYSLYSPHEGLYALATHRSERVALGILYTLERHLEUSERPROVALGIALAGLNR 166
 DB 20416 CGCGCGGACACCGCATCCAGATCACCAGCGCATCCCTGACAGTGGCCAT 20475
 QY 166 GALASNALEUALAALAGLNGLYLEUALASPLYVALSERPHEGLINVALALAS 186
 DB 20476 CGCCCTGATTCGCGACCGGACGCGGACTAGCCACCGGCTCTCTCGTGGTCTGA 20535

OY 332 GIUGLYTTLyLysAspLeuIleIleSphelaIleIleHrcysArgLysPro 349
Db 1198 GGTGGTAGACAAATTTTACTCCAAATGATGTTGTACGTTGTAGAAACCA 1251

RESULT 7

US-08-764-233A-1

Sequence 1, Application US/08764233A
Patent No. 5716849

GENERAL INFORMATION:

APPLICANT: Ligon, James M.
APPLICANT: Schupp, Thomas
APPLICANT: Beck, James J.
APPLICANT: Hill, Dwight S.
APPLICANT: Neif, Snezana
APPLICANT: Ryals, John A.
TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,233A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/729,214
FILING DATE: 09-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: 1506/CIP6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ. ID NO.: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 49377 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Sorangium cellulosum
IMMEDIATE SOURCE:
CLONE: p98/1, pJL3, and pVKM15
FEATURE:

NAME/KEY: misc.feature

LOCATION: 383..760
OTHER INFORMATION: /product- "Sora"
OTHER INFORMATION: /note- "This gene encodes a protein that is highly homologous
OTHER INFORMATION: the reductase domains of type I PKs such as eryA from
OTHER INFORMATION: Saccharopolyspora erythraea."

NAME/KEY: misc.feature

LOCATION: 927..19874
OTHER INFORMATION: /product- "Sora"
OTHER INFORMATION: /note- "Gene product is highly homologous to type I PKs that
OTHER INFORMATION: are known to be involved in the synthesis of polyketide
FEATURE: compounds."
NAME/KEY: misc.feature

LOCATION: 942..7115

OTHER INFORMATION: /product- "Module 1 of Sora"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 7203..12884
OTHER INFORMATION: /product- "Module 2 of Sora"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 13455..19616
OTHER INFORMATION: /product- "Module 3 of Sora"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 19871..46318
OTHER INFORMATION: /product- "Sora"
OTHER INFORMATION: /note- "Gene product is highly homologous to type I PKs 9"

NAME/KEY: misc.feature

LOCATION: 19870..24556
OTHER INFORMATION: /product- "Module 1 of Sora"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 24638..30820
OTHER INFORMATION: /product- "Module 2 of Sora"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 30801..35446
OTHER INFORMATION: /product- "Module 3 of Sora"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 35528..40114
OTHER INFORMATION: /product- "Module 4 of Sora"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 40190..46318
OTHER INFORMATION: /product- "Module 5 of Sora"

NAME/KEY: misc.feature

LOCATION: 46851..47891
OTHER INFORMATION: /product- "Sora"
OTHER INFORMATION: /note- "The protein encoded by the sora gene is highly
OTHER INFORMATION: homologous to the methyltransferase from Streptomyces
OTHER INFORMATION: hygroscopicus that is involved in the synthesis of the
OTHER INFORMATION: polyketide rapamycin."

NAME/KEY: misc.feature

LOCATION: 46851..47891
OTHER INFORMATION: /product- "Sora"
OTHER INFORMATION: /note- "The protein encoded by the sora gene is highly
OTHER INFORMATION: homologous to the methyltransferase from Streptomyces
OTHER INFORMATION: hygroscopicus that is involved in the synthesis of the
OTHER INFORMATION: polyketide rapamycin."

NAME/KEY: misc.feature

LOCATION: 46851..47891
OTHER INFORMATION: /product- "Sora"
OTHER INFORMATION: /note- "The protein encoded by the sora gene is highly
OTHER INFORMATION: homologous to the methyltransferase from Streptomyces
OTHER INFORMATION: hygroscopicus that is involved in the synthesis of the
OTHER INFORMATION: polyketide rapamycin."

NAME/KEY: misc.feature

LOCATION: 46851..47891
OTHER INFORMATION: /product- "Sora"
OTHER INFORMATION: /note- "The protein encoded by the sora gene is highly
OTHER INFORMATION: homologous to the methyltransferase from Streptomyces
OTHER INFORMATION: hygroscopicus that is involved in the synthesis of the
OTHER INFORMATION: polyketide rapamycin."

NAME/KEY: misc.feature

LOCATION: 46851..47891
OTHER INFORMATION: /product- "Sora"
OTHER INFORMATION: /note- "The protein encoded by the sora gene is highly
OTHER INFORMATION: homologous to the methyltransferase from Streptomyces
OTHER INFORMATION: hygroscopicus that is involved in the synthesis of the
OTHER INFORMATION: polyketide rapamycin."

NAME/KEY: misc.feature

LOCATION: 46851..47891
OTHER INFORMATION: /product- "Sora"
OTHER INFORMATION: /note- "The protein encoded by the sora gene is highly
OTHER INFORMATION: homologous to the methyltransferase from Streptomyces
OTHER INFORMATION: hygroscopicus that is involved in the synthesis of the
OTHER INFORMATION: polyketide rapamycin."

NAME/KEY: misc.feature

LOCATION: 46851..47891
OTHER INFORMATION: /product- "Sora"
OTHER INFORMATION: /note- "The protein encoded by the sora gene is highly
OTHER INFORMATION: homologous to the methyltransferase from Streptomyces
OTHER INFORMATION: hygroscopicus that is involved in the synthesis of the
OTHER INFORMATION: polyketide rapamycin."

NAME/KEY: misc.feature

LOCATION: 46851..47891
OTHER INFORMATION: /product- "Sora"
OTHER INFORMATION: /note- "The protein encoded by the sora gene is highly
OTHER INFORMATION: homologous to the methyltransferase from Streptomyces
OTHER INFORMATION: hygroscopicus that is involved in the synthesis of the
OTHER INFORMATION: polyketide rapamycin."

```

Db 47058 GAGGCCGCTCAAGTGTCCGATCGCTTTTCAACAC---TGGGGG-----ATGTTTCAGGAG 47108
OY 92 PheTYrAspSerAspSerThrValSerLeuSerAspHisArgAlaIleValMet 111
Db 47109 GAGGTTTATCGTGAATGTTCTTCAACAGCTGCCAGATTACAGACAGACAGACGCT 47168
OY 112 IleGlnGlu-----SerLeuArgPheAlaSerValSerGluGluArgSer 126
Db 47169 TATTGGACACACTCTACGCTACACAGCTGAGACAGATTCCGGCCGACAGAAATCGG--- 47225
OY 127 LysTTPProLysSerIleValAspValGlyCysGlyIleGlyCysSerArgTyrLeu 146
Db 47226 -----CCACGCAAGATCTTGAGAGCTGCGCTGCACACATGGGAGGCGCTCAACTCTCT 47279
OY 147 AlaLysLysPheGly---AlaPheSerValGlyIleThrLeuSerProValGlnAlaGln 165
Db 47280 TCCCGGATCGAGGCTGCGACACCTTTCGGGCTGCGGCTGACCTGCGACGACAGCGGTGGAC 47339
OY 166 ArgAlaAsnAlaLeuAlaIleAlaGlnGlyLeuAlaAspLysValSerPheGlnValAla 185
Db 47340 ATCGCCACAGCCCGGCTTCGCGAGCCGGG-----TCGTTGACCTACGTTCAAGGC 47390
OY 186 AspAlaLeuGlnGlnProPheSerAspGlyGlnPheAspLeuValTTPSerMetGluSer 205
Db 47391 GACGCGGAGACACCTTCTTCGCGAGCGGGAATTCGACCTGCTCATTCATTCGAGAGC 47450
OY 206 GlyGluHisMetProAspLysAlaLysPheValGlyGluLeuAlaArgValAlaIleAlaPro 225
Db 47451 TCGCACACCTACCTTAATGTCGCCAAATCATTTCTGAACTGGCGGCGGTGCGGACGC 47510
OY 226 GlyAlaIleIleIleIleValThrTTPCysHisArgAspLeuGlyProAspGluGlnSer 245
Db 47511 GCGGGGCTC-----TTTTCGATGTGGATGTCTTCGCGACAAATCGC--- 47552
OY 246 LeuHisProTTPGluGlnAspLeuLeuLysLysIleCysAspAlaTyrTyrLeuProAla 265
Db 47553 -----TATTGGGTGATGCAG 47567
OY 266 TTPCys-----SerThrSerAspTyrValLysLeuGlnSerLeuGlnAsp 283
Db 47568 AATTGACAGCAGACAGACAGCGGCGAGCTGCACTGCTCAAGAGACCGACATCTCGGAG 47627
OY 284 IleLysSerGluAspTTPSerArgPheValAlaIleProPheTTPProAlaValIleArgSer 303
Db 47628 TACGTAAGAAGAGCTATCCGCGACGACGACTGCCCGCCGCG----- 47666
OY 304 AlaPheThrTTPLysGlyLeuSerSerLeuLeuSerSerGlyGlnLysTyrIleLysGly 323
Db 47667 -----AGCAAGCAAGGCAACAGAGTGCAGCGC 47693
OY 324 AlaLeuAlaMetProLeu-----MetIleGluGlyTyrLysLys 336
Db 47694 GCGCTCCCGTACCCGCGCGAGAGCTTTTTCAGCTCGGTGATGAGCGCGCTACGAGATCC 47753
OY 337 AspLeuIleLysPheAlaIle 343
Db 47754 GAG-----TTTGGCGCTC 47765

```

```

; ORGANISM: Escherichia coli
US-08-844-305-1
Alignment Scores:
Pred. No.: 7,18e-09 Length: 1149
Score: 156.50 Matches: 69
Percent Similarity: 41.45% Conservative: 45
Best Local Similarity: 25.09% Mismatches: 105
Query Match: 8.55% Indels: 56
DB: 2 Gaps: 14

US-09-857-613a-28 (1-350) x US-08-844-305-1 (1-1149)
OY 71 GluPheTyrAspGluSerSerGlyLeuTTPGluAsnIleTTPGlyAspHisMetHis 90
Db 352 GACCATTTAGATTTGGGTATATGACTTTGTCAGCCGCGATGTTATCCCTTCATGCAATAT 411
OY 91 Gly-----PheTyrAspSerAspSerThrValSerLeuSerAspHisArgAlaIle 107
Db 412 TCCGCGCTTACTGGAAGATGCCGAT-----AATCTGGAATCTGCCACAGCGC 462
OY 108 GlnIleArgMetIleGlnGluSerLeuArgPheAlaSerValSerGluGluArgSerLys 127
Db 463 AACCTAAATGATTTGTGAAAATTGCACTTA-----AAA 498
OY 128 TTPProLysSerIleValAspValGlyCysGlyIleGlyCysSerArgTyrLeuAla 147
Db 499 CCAAGGATGGCGCTACTGATATTTGGCTGGCGGCGGACATGCACTACATGCA 558
OY 148 LysLysPheGlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaIleAlaGln 167
Db 559 TCTAATATATGACCTAAGCCGCTGGGCGCTCCACCATTTCTCCGACACGCAAAAATGCGT 618
OY 168 AsnAlaLeuAlaIleAlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAla 187
Db 619 CAG-----GACGCTGTGAAGCCTGATGTACCATTT----- 651
OY 188 LeuGlnGlnProPheSerAsp-----GlyGlnPheAspLeuValTTPSerMetGluSer 205
Db 652 TTGCTGCAACATTAATGCTGACCTGACGACGACGAGTTGATGTTATTTCTGTGGGATG 711
OY 206 GlyGluHisMetProAspLys-----AlaLysPheValGlyGluLeuAlaIleAlaGln 223
Db 712 TTCGAGCAGCCTCGACCGAAGAAATTAACATACCTATTTCGCTGGTGATGTAATTTG 771
OY 224 AlaProGlyAlaIleIleIleIleValThrTTPCysHisArgAspLeuGlyProAspGlu 243
Db 772 AAACCGGAGGACATATTCTGCTCCATACT-----ATCGGTTGCAAAAAA 816
OY 244 GlnSerLeuHis-----ProTTPGluGlnAspLeuLysLysIleCysAspAlaTyr 261
Db 817 ACCGATCTGAATGTGATCCCTG-----ATTAAATAATAT 852
OY 262 TyrLeuProAlaTTPCysSerThrSerAspTyrValLysLeuGlnSerLeuSerLeu 281
Db 853 ATTTTTCGAACGTTCTGCGCCCTCT-----CTACGCAAGATGCTCAATCAGCGAA 906
OY 282 GlnAspIleLysSerGluAspTTPSerArgPheValAlaIleProPheTTPProAlaValIle 301
Db 907 CCCACCTTTGTGATGGAAGACTGTCATACCTTCGCTCGTCAATTC----- 951
OY 302 ArgSerAlaPheThrTTPLysGlyLeuSerSerLeuLeuSerSerGlyGlnLys----- 319
Db 952 -----GATACATACGTTGATGCGGTGATGAAACGATTCCTCGCGGATGCGCAAAATTCG 1008
OY 320 -----ThrIleLysGlyAlaLeu-----AlaMetProLeuMet 330
Db 1009 GATAACTATATGTAACGCTTTAAACGAATGTTTACCTATTATC 1051

```

```

RESULT 9
US-09-036-987A-1/c
; Sequence 1, Application US/09036987A
; Patent No. 6143526
; TYPE: DNA

```



```

GENERAL INFORMATION:
APPLICANT: Baltz, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Merlo, Donald J.
APPLICANT: Treadway, Patti J.
APPLICANT: Turner, Jan R.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
TITLE OF INVENTION: Production
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences LLC Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,987A
FILING DATE: 09-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R.
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 80161 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-036-987A-1

Alignment Scores:
Pred. No.: 2e-05 Length: 80161
Score: 152.50 Matches: 48
Percent Similarity: 38.38% Conservative: 28
Best Local Similarity: 24.24% Mismatches: 77
Query Match: 8.33% Indels: 45
Caps: 4

US-09-857-613a-28 (1-350) x US-09-036-987A-1 (1-80161)

OY 85 GLYASPHISMETHISHISGLYPHETRYASP----- 94
Db 13517 GGGCGTCCCTGACACGAGTTACTGGCGCGGATCGGAGAGATCCGCGTCCACA 13458
OY 95 -----SeraspserThrValSerLeuSerAspHisArgAlaAlaGlnIleArgMetIle 112
Db 13457 CCGTGTGGATGCTGGCGACCACTGACGAC-----CTGTTCATCCAGC 13413
OY 113 GlnGluSerLeuArgPheAlaSerValSerGluGluArgSerLysTrpProLysSerIle 132
Db 13412 AAGGCCCGCTCCCTCCCGGAGC-----CACCTG 13383
OY 133 ValAspValGlyCysGlyIleGlyCysSerSerArgTyrLeuAlaLysLysPheGlyAla 152
Db 13382 TTCGACCTGGCTCGGCAATGAGCCGCTAGTCCGTGGCGGACGCGCGCTT 13323
OY 153 ThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAsnAlaLeuAla 172
Db 13322 CGAGTACACCGAATCACCCTGAACGCCAGCATCTCGCGCGCCACCAAGGCTCGCCAAC 13263

```

```

OY 173 AlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAlaLeuGlnGlnProPhe 192
Db 13262 GAGACCGGACTGCGCGACGACTTGTAGTTCAGTCCAGCGCGCCCTGCTCCCTAC 13203
OY 193 SerAspGlnGlnPheAspLeuValTrpSerMetGluSerGlyGlnHisMetProAspLys 212
Db 13202 CCGGACGCTTCTTTCAGCGCCGCGATGGCGATGCTGCTGTCGATGCTGACGAC 13143
OY 213 AlaAspValGlyGlyLeuAlaArgValAlaAlaProGlyAlaIleIleIleVal 232
Db 13142 GCCCGCCGATCCGCGAGCTTCACCGAATCTGGAACCGCGCGCTGCTCTCGGA 13083
OY 233 ThrTrpCysHisArgAspLeuGlyProAspGluGlnSerLeuHisProTrpGluGlnAsp 252
Db 13082 GACATCATCACTCGGCTTCGATCCCGAAGAC----- 13050
OY 253 LeuLeuLysLysIleCysAspAlaTyrTyrLeuProAlaTrpCysSerThrSer 270
Db 13049 -----TACGCGCGGCTTGGACGGCAGCAC 13023

RESULT 10
US-09-370-700-1/c
Sequence 1, Application US/09370700
Patent No. 6274350
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Treadway, Patti J.
APPLICANT: Turner, Jan R.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
FILE REFERENCE: 50489 DIV1
CURRENT APPLICATION NUMBER: US/09/370,700
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/36987
NUMBER OF SEQ ID NOS: 39
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 1
LENGTH: 80161
TYPE: DNA
ORGANISM: Saccharopolyspora spinosa
US-09-370-700-1

Alignment Scores:
Pred. No.: 2e-05 Length: 80161
Score: 152.50 Matches: 48
Percent Similarity: 38.38% Conservative: 28
Best Local Similarity: 24.24% Mismatches: 77
Query Match: 8.33% Indels: 45
Caps: 4

US-09-857-613a-28 (1-350) x US-09-370-700-1 (1-80161)

OY 85 GLYASPHISMETHISHISGLYPHETRYASP----- 94
Db 13517 GGGCGTCCCTGACACGAGTTACTGGCGCGGATCGGAGAGATCCGCGTCCACA 13458
OY 95 -----SeraspserThrValSerLeuSerAspHisArgAlaAlaGlnIleArgMetIle 112
Db 13457 CCGTGTGGATGCTGGCGACCACTGACGAC-----CTGTTCATCCAGC 13413
OY 113 GlnGluSerLeuArgPheAlaSerValSerGluGluArgSerLysTrpProLysSerIle 132
Db 13412 AAGGCCCGCTCCCTCCCGGAGC-----CACCTG 13383
OY 133 ValAspValGlyCysGlyIleGlyCysSerSerArgTyrLeuAlaLysLysPheGlyAla 152
Db 13382 TTCGACCTGGCTCGGCAATGAGCCGCTAGTCCGTGGCGGACGCGCGCTT 13323

```



```

: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961.527
: FILING DATE:
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Brookes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PB340P1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 222:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3236 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: US-08-961-527-222

Alignment Scores:
Pred. No.: 6,82e-06 Length: 3236
Score: 137.50 Matches: 59
Percent Similarity: 40.98% Conservative: 41
Best Local Similarity: 24.18% Mismatches: 85
Query Match: 7.51% Indels: 60
DB: Gaps: 10

US-09-857-613a-28 (1-350) x US-08-961-527-222 (1-3236)
QY 42 SerAlaIaSerSerGluArgGlyGlu-----IlleValLeuGluGlnLysPro 57
Db 2260 TCTAGTTATCAATTAATAGATTAGATTGTCAGATAATATTCATTTGGTTATTAATGAA 2319
QY 58 LysLysAspAspLysLysLysLeuGlnLysGlyIleAlaGluPheTyrAspGluSer 77
Db 2320 GAAAAAAGAGAGGTCTTCAAAATGTCAGAACGACGATAGTTTTCAGCAAAAT----- 2373
QY 78 GlyLeuTPrGluAsnIleTPrGlyAspHisMetHisGlyPheTyrAspSer 97
Db 2374 -----TCGGGAAAAAAGCGTTAC----- 2391
QY 98 ThrValSerLeuSerAspHisArgAlaIleGlnIleArgMetIleGlnIleuSerLeuArg 117
Db 2392 -----GTCACGCTGGAAAGCGTCCACAGATTGG-TTAAATGCAAGAA----- 2432
QY 118 PheAlaSerValSerGluGluArgSerLysTPrProLysSerIleValAspValGlyCys 137
Db 2433 ---GGAGGATTTTCAAAAGAA-----AAGGATTAAGTAGAGGTTGGGTGT 2474
QY 138 GlyIleGlyCysSerSerArgTyrIleuAlaLysLysPheGlyAlaThrSerValGlyTle 157
Db 2475 AATAGGGGAACACCAATTCAGTTGGCACACGCTTTTGGCAAGATAAATCGCTGT 2534
QY 158 ThrLeuSerProValGlnIleGlnArgAlaAsnAlaLeuAlaIleGlnIleuAla 177
Db 2535 GATATGAGTCCTCAAGCTTTAGAGTGAATAAAATCTGCTGAAACGGACGAGGTGGTCT 2594
QY 178 AspLysValSerPheGlnValAlaAspAlaLeuGlnIleuProPheSerAspGlyGlnPhe 197
Db 2595 CATTTAATCAGTTTTCAGAGCAATGCAATGCAAACTTCCTATCAGATGCTAGTTT 2654
QY 198 AspLeuVal-----TPrSerMetClnSerGlyLuhMetProAspLys 212
Db 2655 GATATGTTAATAATGACGATATGCTGACTATGCAAGCCGATCAAA-----GCTAAG 2705
QY 213 AlaLysPheValGlyIleuAlaIleArgValAlaAlaProGlyAlaIleIleIle----- 230
Db 2706 AAAAAATGTATATGGAATATCTAAGGGTATTTAAAACTGTGAGGCTCTTCTTGACACAT 2765
```

```

QY 231 -----IlleValThrTPrCysHis 236
Db 2766 GATGCTCTTCTTAAGAGACTAAAGAGCTATACAGACAGGAATTAACAGCAATTCAT 2825
QY 237 ArgAspLeuGlyProAspGluGlnSerLeuHisProTPrGlnIleAspLeuLysLys 256
Db 2826 GTAATGTAGTCTTAACTCAAGAT-----GGTTGGACAGAGTGATGATGATCA 2879
QY 257 ---IlleCysAsp 259
Db 2880 CGTTATTTGTAT 2891

RESULT 13
US-08-576-626A-2
: Sequence 2, Application US/08576626A
: Patent No. 5998194
: GENERAL INFORMATION:
: APPLICANT: Summers, R.G.
: APPLICANT: Katz, L.
: APPLICANT: Donadio, S.
: APPLICANT: Slaver, M.J.
: TITLE OF INVENTION: POLYMERIDE-ASSOCIATED SUGAR
: NUMBER OF SEQUENCES: 60
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Abbott Laboratories
: STREET: 100 Abbott Park Road
: CITY: Abbott Park
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60064-3500
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/576,626A
: FILING DATE: 21-DEC-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Dianne Casuto
: REGISTRATION NUMBER: P-40,943
: REFERENCE/DOCKET NUMBER: 5857.US.01
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (847) 938-3137
: TELEFAX: (847) 938-2623
: TELEX:
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 8051 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-576-626A-2

Alignment Scores:
Pred. No.: 0.000175 Length: 8051
Score: 131.00 Matches: 80
Percent Similarity: 33.23% Conservative: 32
Best Local Similarity: 23.74% Mismatches: 128
Query Match: 7.16% Indels: 97
DB: Gaps: 19

US-09-857-613a-28 (1-350) x US-08-576-626A-2 (1-8051)
QY 18 PheArgSerClnSerProArg-----Thr 25
Db 2153 TGGGAGCGCCCTCACAGAGTGTGAGACACCCACCTACACCGAAGCGGTGAAGC 2212
```

Oy	26	PheIleArgIleIleVal	-----	GLyProArgSerTAlaPro	38	
Db	2213	TTGCGCAGGAGGGCGCTGTCTCCAGCCGACCCCGCAGAGATGCTCCCGGACATGGAGAAC			2272	
Oy	39	IleArgAlaSerAlaAlaSerSerGluArgGlyGluIleValIleGluIleGlnIleLysProLys			58	
Db	2273	TCACGGCGCCGCCACCGCGGTACCGGTTTCCGACCGACAGTCCGTCCGACAGCACACT			2332	
Oy	59	LysAspLysLysLysLysLysIleGlyIleIleAlaGluPheTyrAspLysSerGly			78	
Db	2333	CCGAGCGAGCAGCAGGATGTACGAGGGCGGGTTCGCCGACCTTACGACCG			2383	
Oy	79	LeuIrrpLysnIleIrrpGlyAspHisMetHisIleGlyPheTyrAspSerAspSerThr			98	
Db	2384	-----	TTTACC	CGCGCGCGGGCG	2401	
Oy	99	ValSerLeuSerAspHis	-----	ArgAlaAlaGlnIle	ArgMetIleGlnIleSer	115
Db	2402	-----	AAGACATACGGCGCGCGCGGACCGCGGACAGTCCGCGCTGTCCAGAGACCG		2452	
Oy	116	LeuArgPheAlaSerValSerGluGluIleArgSerLysTrrProLysSerIleValAspAla			135	
Db	2453	CTGGCCCTCGGCTTCC		TTCCCTGCTGCACGTG	2482	
Oy	136	GlyLysGlyIleGlyGlyLysSerArgTyrLeuAlaLysLysPheGlyAlaIleHisSerVal			155	
Db	2483	GCCTTCGGGACCGGACCCACCCACTCGCCCGTTCCCGGACCTTTTCGACAGACGTGAC			2539	
Oy	156	GlyIleThrLeuSerProValGlnAlaGlnArgAla			171	
Db	2540	GGGCTGGAGCTGTGGCGGAGATGACGAGTGCACCGCGCCGACACTCGCGCGCATCCCG			2599	
Oy	172	AlaAlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAlaSerPalaeuGlnIlePro			191	
Db	2600	GTGCTGACGGC	-----	GACATGCGCGACTTCGCGCTG	2632	
Oy	192	PheSerAspGlyGlnPheAspPheValTrrSerMetGlnSerGly			210	
Db	2633	-----	GATCGCGAGTTTCGACCGCCGTCACCTGATTTTACGTTCCATCGGCGACATCGCGC		2686	
Oy	211	AspLysAlaLysPhe			227	
Db	2687	GACGGCGCGGACGTGAGACAGAGCGGTGCGCTTCTGCGCGGACACTCGCGCGCGCGCGCG			2746	
Oy	228	IleIleIleIleValIleThrTrrCysHisArgAspLeuGlyProAspGlnIleSerLeuHis			247	
Db	2747	GTGCGGTGTGTGCAACCGTG			2767	
Oy	248	ProTrrPheGluGlnAspLeuLysLysIleCysAspAlaTyrTyrLeuProAlaTrr			266	
Db	2768	---TGTGTTCCCGGAGCACTTCTC		---GAGCGCTACCTGTGGCGGTGACGTG	2812	
Oy	266	PCysSerThrSerAspTrrValLysLeuLeuGlnSerLeuSerLeuGlnAspIleLysSe			286	
Db	2813	GTGGCGGACGGCAGACTGACAGATCTCGCGGTCTCGCAGCTCGTGGCGCGCGCGCGCG			2872	
Oy	286	TGluAspTrrSerArgPheValAlaProPheTrrProAla			304	
Db	2873	ACCGGAGTGAAGATCC	-----	ACTGGGTGTGGCGGACGCGGTGAACGCTTCGCGCGCAC	2926	
Oy	304	AlpheIrrTrrLysGlyLeuSerSerIleLeuSerSerGlyGlnLysThr			320	
Db	2927	C---ACGTGGAGCACTACAGATATTCACCTCTTGTGACGGAGCAGCATACG			2972	

RESULT 14
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.

```

: APPLICANT VENTER, John C.
: TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
: TITLE OF INVENTION: TUBERCULOSIS
: FILE REFERENCE: 24366-20007.00
: CURRENT APPLICATION NUMBER: US/09/103,840A
: CURRENT FILING DATE: 1998-06-24
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2
: LENGTH: 4403765
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: FEATURE:
: OTHER INFORMATION: CDC 1551
: OTHER INFORMATION: "n" bases at various positions throughout the sequence
: OTHER INFORMATION: represent a, t, c or g
: US-09-103-840A-2

```

Alignment Scores:	
Pred. No.:	13.5
Score:	127.00
Percent Similarity:	40.37%
Best Local Similarity:	25.23%
Query Match:	6.94%
DB:	4
	Gaps: 5

```

OY 95 SerAspSerThrValSerLeuSerSerPheHisArgAlaAlaGlnIleArgMetIleGlnIu 114
Db 4172057 GCCCATGGACATGCCACACAGCAGATCAAAAGCGTGTGGCCGACGCTATTCCCGGGA 41721116
OY 115 SerLeuArgPheAlaSer----- 120
Db 4172117 CATGATGCGCCTTGCTACCGGTAGCTCCTTTGACCCGGCGCGGACGACATTGACCCGTAG 4172176
OY 121 -----ValSerGluGlnArgSerIuSTTProIySerIleValAspValGly 136
Db 4172177 TTGGCTGACCAACATCGGGGCTGAGCTGCACAGCGACCGCGCGCGGTGCGCGACATGCGCC 4172233
OY 137 CysGlyIleGlyGlySerSerArgTyrLeuAlaLysLysPheGlyAlaThrSerValGly 156
Db 4172237 GCCCGGGCGCGCGCGCTCGGCACGGGCTGTGGCCAGCGAGCATAGGATGCTGCTGCAGCGG 4172299
OY 157 IlePheLeuSerProValGlnAlaIleGlnArgAlaAspAlaLeuAlaAlaIleGlnIyLeu 176
Db 4172297 GTGCACATACACGAGATGACGTGAAGCGCGCCGACAGCGCGCGTGGCCAAACCGCGCTG 4172356
OY 177 AlaAspLysValSerPheGlnIleValAlaAspAlaLeuGlnIleProPheSerAspGlnI 196
Db 4172357 ACCGAGCGGGGTGCGCTTCCACCTGGCGACGCGCGCGAGTACGTCCGCTTCCCGGACGACGA 4172411
OY 197 PheAspLeuValTyrSerMetGlnSerGluGlnIleHisMetProAspLysAlaLysPheVal 216
Db 4172417 TTGCAGCGCGCTGAGTGTGCGAGTGGCGCGCTTGCACATTCCCGGACAGAACAGCGCGCGCC 4172476
OY 217 GlyGluIleAlaAlaArgValAlaAlaIleProGluAlaIle-----IleIleIleValThrP 234
Db 4172477 CAGCATGTCGTGCGGATATCTCGCTCTGTGTCGTGCGCGCGCGGACATCCGAGATGTACAGTCTC 4172530
OY 235 CysHisArgAspLeuGluIleProAspGluGlnIleSerLeuHisProTyrGluGlnAspLeuLeu 254
Db 4172537 GGGAGAGGGCGGCGCTCGCGGAGAGTGAACCCGATGGCGCGGTGG-----GTG 4172584
OY 255 LysLysIleCysAspAlaIleTyrTyrLeuProAlaIleTyrGlySerThrSerAspTyrValLys 274
Db 4172585 GCCTGCATCGCGCGAGCGCGCA-----ACCGTACCGAGTACACCGGAC 4172620
OY 275 LeuIleGlnSerLeuSerLeu-----GlnAspIleLysSerGlnAspThrSer 290
Db 4172627 ATCTCGAAGGGCGCGGATTCGCGACCGCGGACATCGAGTGTGCATGACGAGAGC 4172680

```


GenCore version 5.1.4-P5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 19, 2003, 08:06:01 ; Search time 88 Seconds
(without alignments) updates/sec
3084.279 Million cell

Title: US-09-857-613A-28

Perfect score: 1830
Sequence: 1 MATVVRIPRTISCHIHIFPRS.....IEGKKDLIKFAIITCRKE 350

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 538826 seqs, 387737923 residues

Total number of hits satisfying chosen parameters: 1077652

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09857613/rnunat.12032003.100038.6130/app-query.fasta.1.519
-DB=PublishedApplications_NA -OUT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPT=0 -LOOPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUT=fastap -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09857613 -QCN_1.1.53 -rnunat.12032003.100038.6130
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGOQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database: Published Applications_NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCIT_NEM_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description
1	386.5	21.1	252	10	US-09-878-574-695	Sequence 695, App
2	316	17.3	311	10	US-09-294-093B-3472	Sequence 3472, AP
3	226	12.3	1080	9	US-09-938-842A-1815	Sequence 1815, AP
4	226	12.3	1497	10	US-09-779-144A-6	Sequence 6, App11

5	212.5	11.6	1086	9	US-09-938-842A-1856	Sequence 1856, Ap
6	203.5	11.1	1152	10	US-09-801-368-95	Sequence 95, Appl
7	203.5	11.1	1320	10	US-09-779-144A-1	Sequence 1, Appl1
8	203.5	11.1	1320	10	US-09-779-144A-5	Sequence 5, Appl1
9	196.5	10.7	1420	10	US-09-779-144A-3	Sequence 3, Appl1
10	156.5	8.6	3900	10	US-09-805-681-1	Sequence 1, Appl1
11	148.5	8.1	411	10	US-09-878-574-3629	Sequence 3629, Ap
12	141	7.7	753	9	US-10-260-877-53	Sequence 53, Appl1
13	140	7.7	1801	10	US-09-845-248-1	Sequence 1, Appl1
14	139	7.6	1353	10	US-09-738-626-636	Sequence 636, Ap
15	128	7.0	290	10	US-09-294-093B-5022	Sequence 5022, Ap
16	123	6.7	891	9	US-09-938-842A-1172	Sequence 1172, Ap
17	117.5	6.4	618	9	US-09-738-626-1423	Sequence 1423, Ap
18	114	6.2	761	10	US-09-938-626-984	Sequence 984, Ap
19	113	6.2	771	10	US-09-815-242-7986	Sequence 7986, Ap
20	113	6.1	4315	9	US-09-070-927A-1	Sequence 1, Appl1
21	112	6.1	753	10	US-09-738-626-2865	Sequence 2865, Ap
22	112	6.1	3309400	9	US-09-738-626-1	Sequence 1, Appl1
23	110	6.0	654	9	US-09-738-626-1499	Sequence 1499, Ap
24	110	6.0	969	9	US-09-938-842A-329	Sequence 329, Ap
25	104	5.7	5793	10	US-09-880-107-2109	Sequence 2109, Ap
26	101	5.5	699	10	US-09-974-300-1012	Sequence 1012, Ap
27	101	5.5	1406	10	US-09-939-980-119	Sequence 119, Ap
28	99	5.4	726	10	US-09-815-242-8504	Sequence 8504, Ap
29	99	5.4	30365	10	US-09-825-414-1	Sequence 1, Appl1
30	98.5	5.4	702	10	US-09-815-242-4259	Sequence 4259, Ap
31	97.5	5.3	2313	9	US-09-938-842A-1580	Sequence 1580, Ap
32	97	5.3	801	12	US-10-007-693-50	Sequence 50, Appl
33	97	5.3	3309400	9	US-09-738-626-1	Sequence 1, Appl1
34	95.5	5.2	1836	9	US-09-938-842A-651	Sequence 651, App
35	95.5	5.2	3198	9	US-10-152-661-601	Sequence 601, App
36	95.5	5.2	3198	9	US-09-866-050A-601	Sequence 601, App
37	94.5	5.2	2040	9	US-09-738-626-1377	Sequence 1377, Ap
38	94.5	5.2	2340	10	US-09-738-626-1377	Sequence 7995, Ap
39	94	5.1	7596	10	US-09-815-242-7995	Sequence 1, Appl1
40	94	5.1	1500	9	US-09-728-952-1	Sequence 1281, Ap
41	93	5.1	2811	9	US-09-938-842A-1281	Sequence 2538, Ap
42	92.5	5.1	2811	9	US-09-938-842A-2538	Sequence 8247, Ap
43	92	5.0	2670	10	US-09-878-574-8247	Sequence 27, Appl
44	92	5.0	23907	9	US-09-927-627-27	Sequence 6, Appl1
45	92	5.0	24120	9	US-10-077-130-6	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-09-878-574-695
Sequence 695, Application US/09878574
Patent No. US20020110548A1
GENERAL INFORMATION:
APPLICANT: Byrum, Joseph R.
APPLICANT: La Rosa, Thomas J.
APPLICANT: Thompson, Michael D.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(13401)B
CURRENT APPLICATION NUMBER: US/09/878,574
PRIOR FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 09/333,535
NUMBER OF SEQ ID NOS: 15775
SEQ ID NO 695
LENGTH: 252
TYPE: DNA
ORGANISM: Glycine max
OTHER INFORMATION: Clone ID: LIB3028-047-Q1-B1-D12
US-09-878-574-695

Alignment Scores:
Pred. No.: 1.93e-36
Score: 386.50
Percent Similarity: 93.83%
Best Local Similarity: 90.12%
Length: 252
Matches: 73
Conservative: 3
Mismatch: 4


```

Db 328 ATCCAGCAAGAAATG-----GCCGTGATCTCATCAAGTGAACCGGGCAAAAG 378
Qy 132 ILeValAspValGlyCysGlyIleGlySerSerArgTyrLeuAlaLysLysPheGly 151
    |||:::||||| |||:::|||||
Db 379 ATTCTGACCGCTGCTGGCGGTGGGCGCATGAGACCATCGGGCCCATTCAG 438
Qy 152 AlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnAlaAsnAlaLeuAla 171
    |||:::||||| |||:::|||||
Db 439 GCCCAAGTCACTGCAATCATCATCAACAGTACCAAGTGAACGCAAGCCTTCACAAAC 498
Qy 172 AlaAlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAlaGlnGlnPro 191
    |||:::||||| |||:::|||||
Db 499 AAGAAAGCTGAGCTGATCTCTCTGCAACGCTGCTTGTGCTTAACTTTAAAGATCCG 558
Qy 192 PheSerAspGlyGlnPheAspLeuValTyrSerMetGluSerGlyLysHisMetProAsp 211
    |||:::||||| |||:::|||||
Db 559 TTGATGATAAAGACGCTTGAAGCGGCTTACTCGATGAGAGCTGCTGCTGCTGCTGCT 618
Qy 212 LysAlaLysPheValGlyGlnLeuAlaArgValAlaAlaProGlyAlaIleIleIle 231
    |||:::||||| |||:::|||||
Db 619 CTCGAAGAACTATATCTGCGAGATCTTCAGATGATGAACACGAGATCTTGTGCTGCC 678
Qy 232 ValThrProCysHisArgAspLeuGlyProAspGlnGlnSerLeuHisProTyrGlnGln 251
    |||:::||||| |||:::|||||
Db 679 TACCAATGGGTCCACCTGAAATAATACAGACGATGACGAAGAACAC-----AAG 729
Qy 252 AspLeuLeuLysLysIleCysAspAlaTyrTyrLeuProAlaTyrCysSerThrSerAsp 271
    |||:::||||| |||:::|||||
Db 730 GAGGTGATTCAGAGGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 789
Qy 272 TyrValLysLeuLeuGlnSerLeuSerLeuGlnAspIleLysSerGluAspTyrSerArg 291
    |||:::||||| |||:::|||||
Db 790 ATGAGCCCTGACGCGGAAGAGTGGGCTTGAAGTGAAGAGAGAGAGAGAGAGAGAGAG 849
Qy 292 PheValAla---ProPheTyrProAlaValIle-----ArgSerAlaPheThrTyrLys 308
    |||:::||||| |||:::|||||
Db 850 CCACCGCTAAACCGGTGGAACCGGTAAAGATGGAAGATGCTTAT---TGAGAGA 906
Qy 309 Gly-----LeuSerSerLeuLeuSerSerGlyGlnLysThrIleLysGlyAlaLeuAla 326
    |||:::||||| |||:::|||||
Db 907 AACCATTTGCTGCTGATCTTCTTCTGCTATTTGGGTGCTGCTTAAAGAGACTGTGAT 966
Qy 327 MetProLeuMetIle-----GlnGlyTyrLysLysAspLeu 338
    |||:::||||| |||:::|||||
Db 967 GTTCATATAGATGTTGTTAAGACTGCTGATATTATGACCAAGAGCTGGAGAGACTGATC 1026
Qy 339 IleLys---PheAlaIleIleThrCysArgLysProGlu 350
    |||:::||||| |||:::|||||
Db 1027 TTCTCTCCGATGATATGATTCTCTGTAGAAACCAAGAG 1065

```

RESULT 4

```

US-09-779-144A-6
; Sequence 6, Application US/09779144A
; Patent No. US20020148006A1
; GENERAL INFORMATION:
; APPLICANT: Nes, David W.
; TITLE OF INVENTION: Transgenic Plants with Modified Sterol Compositions
; FILE REFERENCE: 11899.0198 DVS01 MOBF.198
; CURRENT APPLICATION NUMBER: US/09/779,144A
; PRIORITY FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: US 09/106,926
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Zea mays
US-09-779-144A-6

```

Alignment Scores:

```

Pred. No.: 1,47e-16 Length: 1497
Score: 226.00 Matches: 81

```

```

Percent Similarity: 38.77% Conservative: 45
Best Local Similarity: 24.92% Mismatches: 121
Query Match: 12.35% Indels: 78
DB: 10 Gaps: 10

```

US-09-857-613a-28 (1-350) x US-09-779-144A-6 (1-1497)

```

Qy 69 ILeAlaGluPheTyrAspGluSerSerGlyLeuTyrGlnAlaIleTyrGlyAspHisMet 88
    |||:::||||| |||:::|||||
Db 216 GTTATATAATATATATATATATATATATATATATATATATATATATATATATATATATAT 275
Qy 89 HisIleGlyPheTyrAspSerAspSerThrValSerLeuSerAspHisArgAlaAlaGln 108
    |||:::||||| |||:::|||||
Db 276 CACTTTCCTCAAGATGGAAGAGAAATCCCTTACGTAACCAATCAACGCAATGAGCAT 335
Qy 109 ILeArgMetIleGlnGlnSerLeuArgPheAlaSerValSerGlnGlnArgSerLysTyr 128
    |||:::||||| |||:::|||||
Db 336 TTCTTGGCCCTGCAACTGTTGTTG-----AAACCA 365
Qy 129 ProLysSerIleValAspValGlyCysGlyIleGlyLysSerArgTyrLeuAlaLys 148
    |||:::||||| |||:::|||||
Db 366 GCAATGAAAGCTTTTATAGATGTGCTGTGATGATGATGATGATGATGATGATGATGATGAT 422
Qy 149 LysPheGlyAlaThrSerVal---GlyIleThrLeuSerProValGlnAlaGlnAla 167
    |||:::||||| |||:::|||||
Db 423 AGATTAGCTCAACTTCACTTACCGGATTAATCAACGAATACCAATACCAAGAGGGA 482
Qy 168 AsnAlaAlaAlaAlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAlaAla 187
    |||:::||||| |||:::|||||
Db 483 AAGAGCTCAACCGTTTACGAGAAATTAATGCAACATGATGATTTGTCAAGCGGACTTC 542
Qy 188 LeuGlnGlnProPheSerAspGlyGlnPheAspLeuValTyrSerMetGluSerGlyLys 207
    |||:::||||| |||:::|||||
Db 543 ATGAAGATGCCCTTGCATGACACACCTTTGATGCTGTTTACGCAATGAGCAACATGCT 602
Qy 208 HisMetProAspLysAlaLysPheValGlyGlnLeuAlaArgValAlaAlaProGlyAla 227
    |||:::||||| |||:::|||||
Db 603 CATCACCCTGATCCAGTGTGTTGCTACAGAGATATATCGTGTGTTGAAGCTGGCCAG 662
Qy 228 IleIleIleIleValAlaThrTyrCys-----His 236
    |||:::||||| |||:::|||||
Db 663 TGCCTTGGCGGTGACAGTGGTGGCTTACGATCACTATGATCTCAACATGCAACCCAC 722
Qy 237 Arg-----AspLeuGly-----ProAspGlnGlnSerLeuHis 247
    |||:::||||| |||:::|||||
Db 723 AAAAGATATCAAGATGAAATTAATGAGCTTGGCAATGCGCTGCAGATATCAAGACCTGG 782
Qy 248 Pro-----TyrGlnGlnAspLeu 254
    |||:::||||| |||:::|||||
Db 783 CAATGCTCCGCGAGTAAAGACCGCGGTTTGAGGTTGTTGGATGATGATGATGATGAT 842
Qy 255 LysLysIleCysAspAlaTyrTyrLeuProAlaTyrCysSer----- 268
    |||:::||||| |||:::|||||
Db 843 GAAATATCTCCCTTGCCTTGTGATCTGCTTGCATTCACACCGCATTCCTCCCTGATGAC 902
Qy 269 -----ThrSerAspTyrValLysLeuGlnSer 278
    |||:::||||| |||:::|||||
Db 903 TTCCGTTTGGACCTCTGTGCGAGCATGATTTACCCGCAAAATGCTAAAGCCCTGAGATAC 962
Qy 279 LeuSerLeuGlnAspIleLysSerGlnAspTyrSerArgPheValAlaProPheThrPro 298
    |||:::||||| |||:::|||||
Db 963 GTTGCTCTGCTCGGAGGAGGAGAGAGGCTCTACTTTCCTGAG----- 1010
Qy 299 AlaValIleArgSerAlaPheThrTyrLysGlyLeuSerSerLeuSerSerGlyLys 318
    |||:::||||| |||:::|||||
Db 1011 -----AAGCTGCAAGAGGCTGTGAGAGGCGGAAG 1043
Qy 319 LysThrIleLysGlyAlaLeuAlaMetProLeuMetIleGlnGlyTyrLysLysAspLeu 338
    |||:::||||| |||:::|||||
Db 1044 AAGGAGATC-----TTCAAGCCCATATGACTCTTTTGTGTCGGAAGCCTCTT 1091
Qy 339 IleLysPheAlaIle 343
    |||:::|||||

```

Db 1092 CTGGAATGAGCTCTT 1106

RESULT 5

US-09-938-842A-1856

Sequence 1856, Application US/09938842A

Patent No. US20020160378A1

GENERAL INFORMATION:

APPLICANT: Harper, Jeff

APPLICANT: Krieps, Joel

APPLICANT: Wang, Xun

APPLICANT: Zhu, Tong

TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

FILE OF INVENTION: SAME, AND METHODS OF USE

CURRENT APPLICATION NUMBER: US/09/938,842A

PRIOR FILING DATE: 2001-08-24

PRIOR APPLICATION NUMBER: US 60/227,866

PRIOR FILING DATE: 2000-08-24

PRIOR APPLICATION NUMBER: US 60/264,647

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/300,111

NUMBER OF SEQ ID NOS: 5379

SEQ ID NO 1856

LENGTH: 1086

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-09-938-842A-1856

Alignment Scores:

Pred. No.: 3,35e-15 Length: 1086
Score: 212.50 Matches: 80
Percent Similarity: 42.54% Conservative: 54
Best Local Similarity: 25.40% Mismatches: 144
Query Match: 11.61% Gaps: 37
DB: 9 Indels: 11

US-09-857-613A-28 (1-350) x US-09-938-842A-1856 (1-1086)

QY 55 GlnLysProLysLys---AspAspLysLysLysLeuGlnLysGlyLeuAlaGluPheTyr 73
Db 175 CGCGCTGCTCAAGAAATGCAAGACCGCGGAGAGTTCCACAGCTTCGCGACACATTTCAC 234
QY 74 AspGluSerSerGlyLeuTrpGluAsnLleTrpGlyAspHisMetHisLysGlyPheTyr 93
Db 235 AATCTGCTCAACCCACATATACAGTGGGATGGGACATCTCCAC-----TCTCTCA 288
QY 94 AspSerAspSerThrValSerLeuSerAsp-----HisArgAlaAlaGlnLle 109
Db 289 CCATCATTCCTCCGGAATATCTCACAAAGACCGCGCTCTCCACGAAGANTGCGCGTA 348
QY 110 ArgMetIleGlnGluSerLeuArgPheAlaSerValSerGluGluArgSerLysTrpPro 129
Db 349 GATCTGATCCAA-----GTCAACCTGCGT 372
QY 130 LysSerIleValAspValGlyCysGlyIleGlyLysSerArgTyrLeuAlaLysLys 149
Db 373 CAAAGATCTTACAGCTGCGTGGATGCGGTCCGATCCGACGATTCGATCTCAC 432
QY 150 PheGlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnAlaAsnAla 169
Db 433 TCGCGAGCTAACGTACGATGCGGATTAACAACGATACAGTGAACAGAGCTGCTC 492
QY 170 LeuAlaAlaAlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAlaLysGln 189
Db 493 CACATTAAGAAAGCTGCTGCGACGCGCTTTCGAGGCTGCTGTGTAACCTCTCCAG 552
QY 190 GlnProPheSerAspGlyGlnPheAspLeuValTrpSerMetGluSerGlyGluHisMet 209
Db 553 ATGCGCTGATGATGACAAAGTTCGACGCTTATTCATGCAAGCCACGCTGACCGG 612
QY 210 ProAspLysAlaLysPheValGlyLeuAlaArgValAlaAlaProGlyValAlaIle 229

Db 613 CCGAAGCTGAGAAGATGTACGACAGATCTACAGGGTGTGAAACCGGATCTATGAT 672
QY 230 IleIleValIleTrpCysHisArgAspLeuLysProAspGluGlnSerLeuHisProTrp 249
Db 673 GTGTCTGACGATGGGTTCAGACGAGAAATTTAAGCGGAGATGACGAACAC----- 726
QY 250 GluGlnAspLeuLeuLysLysIleCysAspAlaIleTyrTrpLeuProIleTrpCysSerThr 269
Db 727 ---GTGAGAGTATCCAGAGGATTCAGAGCGCATCGCTTACCGAGGGCTTAGGCTTAC 783
QY 270 SerAspTyrValLysLeuLeuGlnSerLeuSerLeuGlnAspIleLysSerGlnAspTrp 289
Db 784 GTGATATATAGCTGAGACGCGCTTAAAGAGTTGCGTTTGAATGATGAGAGATGCTG 843
QY 290 SerArgPheValAla---ProPheTrpProAlaValIle-----ArgSerAlaPheThr 306
Db 844 GCGAGTCCACCGCTGACCGCGGTGACACTAGCGCTTAAGATGCGTGAAGGCTTAT 900
QY 307 TrpLysGly-----LeuSerSerLeuLeuSerSerGlyGlnLysThrIleLysGlyAla 324
Db 901 TGGAGGATACATGTTGTGTTCAGATTTTGTACGCGTTGCGAGTTGCTCTTAAGAACT 960
QY 325 LeuAlaMetProLeuMetIle-----GluGlyTyrLysLys 336
Db 961 GTTGATGTTCAATGAGATGTTGTTTAAAGACTCGTATATTGACAGAGAGTGAAC 1020
QY 337 AspLeuIleLys---PheAlaIleIleTrpCysArgLysProGlu 350
Db 1021 GGAATATCTCTCCGATGCATATGATTCCTGCAAAACCGAG 1065

RESULT 6

US-09-801-368-95

Sequence 95, Application US/09801368

Patent No. US20020128250A1

GENERAL INFORMATION:

APPLICANT: Busby, Robert

APPLICANT: Calli, Brian

APPLICANT: Hecht, Peter

APPLICANT: Holtzman, Doug

APPLICANT: Madden, Kevin

APPLICANT: Milne, Todd

APPLICANT: No. US20020128250A1man, Thea

APPLICANT: Royer, John

APPLICANT: Salama, Sofie

APPLICANT: Sherman, Amir

APPLICANT: Silva, Jeff

APPLICANT: Summers, Eric

TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi

FILE REFERENCE: 109272.147

CURRENT APPLICATION NUMBER: US/09/801,368

PRIOR FILING DATE: 2001-03-07

PRIOR APPLICATION NUMBER: US 09/487,558

PRIOR FILING DATE: 2000-01-19

PRIOR APPLICATION NUMBER: US 60/160,587

PRIOR FILING DATE: 1999-10-20

NUMBER OF SEQ ID NOS: 440

SOFTWARE: PatentIn version 3.0

SEQ ID NO 95

LENGTH: 1152

TYPE: DNA

ORGANISM: Saccharomyces cerevisiae

US-09-801-368-95

Alignment Scores:
Pred. No.: 4.09e-14 Length: 1152
Score: 203.50 Matches: 78
Percent Similarity: 37.99% Conservative: 58
Best Local Similarity: 21.79% Mismatches: 149
Query Match: 11.12% Indels: 73
DB: 10 Gaps: 12

US-09-857-613A-28 (1-350) x US-09-801-368-95 (1-1152)

```

Qy 42 SerialaaserSerGluArgGlyLeuIleValleuGluGlnLysProLysLysAsp 61
Db 106 AACTCTGCCCCAAAGAAAGCCGTTGACAGAACTTGGAGATGATGAGAACGAT 165
Qy 62 LysLysLysLeuGlnLysGlyIleAlaGluPheTyrAspGluSerSerGlyLeuTyrp 81
Db 166 AAGATGCGCCAAAGAACGTCGCTT---GAGATATATATGAAGCCACATTCCTACTAT 222
Qy 82 AsnIle-----TrrpGlyAspHisMetHisGly---Phe 92
Db 223 AACCTCGTACAGATTTCTATGAATATGTTGGGCTTCCTTCCATTCAGAGATTT 282
Qy 93 TyrAspSerAspSer-----ThyValSerLeuSerAspHisArgAlaGlnIle 109
Db 283 TATTAAGGTGAGAGATTTGCTGCTCCCTGCTACAGACATGACATTTATTTAGCTTACAG 342
Qy 110 ArgMetIleGlnGluSerLeuArgPheAlaSerValSerGluGluArgSerLysTyrPro 129
Db 343 GCTGATTCACAAAGAGC----- 360
Qy 130 LysSerIleValAspValGlyCysGlyIleGlyLysSerArgTyrLeuAlaLysLys 149
Db 361 GATTAGTTCGACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Qy 150 PheGlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAsnAla 169
Db 421 ACCGCTGTACAGTCATGCTGCTTAAACATTAACATTCCTTCCAAATTCACAAAGCAAAATAT 480
Qy 170 LeuAlaAlaIleGlnGlyLeuAlaAspLysValSerPheGlnValAlaAlaLeuGln 189
Db 481 TACCTTAAATTAACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Qy 190 GlnProPheSerAspGlyGlnPheAspLeuValTyrPserMetGluSerGlyLeuHisMet 209
Db 541 ATGATTTTCGAGAAACACTTTCGACAAAGTTTATGCAATTTGACAGCCATGTCACCT 600
Qy 210 ProAspLysAlaLysPheValGlyLeuLeuAlaArgValAlaAlaProGlyAlaIleIle 229
Db 601 CCAAAATTAAGAGGCTATACAGCAAAATCTACAGATTTTGAACCGGCTGATCCTT 660
Qy 230 IleIleValIleThrPyrCysHisArgAspLeuGlyProAspGluGlnSerLeuHisProTyr 249
Db 661 GCTGTTTACGAATGGGATATGACTGATTAATATGACGAAACATCTGCAACAT----- 714
Qy 250 GluGlnAspLeuLeuLysIleCysAspAlaTyrTyr-----Leu 263
Db 715 -----AGAAAGATC-----GCTTATGAATTTGAACCTAGTGTATGATC 753
Qy 264 ProAlaIlePyrCysSerThrSerAspTyrValLysLeuLeuGlnSerLeuSerLeuGlnAsp 283
Db 754 CCAAAAGATTTCCATGATGACGCTGCTGACGCTGCTGACGCTGCTGCTGCTGCTGCTGCT 813
Qy 284 IleLysSerGluAsp-----TyrPserArgPheValAla 294
Db 814 CTCGTTACGAAAGACTGCGCGGACATGATGATGAATCCCTGATATTAACCATTAAT 873
Qy 295 ProPheTrrpProAlaVal-----IleArgSerAlaPhe 305
Db 874 GGTGAGTGAAGTACGCTCAAAACTTACGTAATTTGCCACATTTTTCAGAACTTCTTAC 933
Qy 306 ThrTrrpLysGlyLeuSerSer-----LeuLeuSer 315
Db 934 TTGGGTACGACATTTTACTACGCAATGCTTACTGTAATGGAATAATTTAGCTTACGCCCA 993
Qy 316 SerGlyGlnLysThrIleLysGlyAlaLeu-----AlaMetProLeuMetIleGln 332
Db 994 GAAAGTTCACAAAGAGATTAATGCTGCTGACGAAATGCGCGGTTGGTTTACTTGGCGGT 1053
Qy 333 GlyTyrLysLysAspLeuIleLysPheAlaIleIleThrCysArgLysProGlu 350
Db 1054 GCTAAGTCCAAAGTATTATTCACCTCAATGATGATTTTCGTCGTAGAGAGCCAGAA 1107

```

```

RESULT 7
US-09-779-144A-1
; Sequence 1, Application US/0979144A
; Patent No. US20020148006A1
; GENERAL INFORMATION:
; APPLICANT: Nes, David W.
; TITLE OF INVENTION: Transgenic Plants with Modified Sterol Compositions
; FILE REFERENCE: 11899,0198, DUS01 MORT:198
; CURRENT APPLICATION NUMBER: US/09/779,144A
; PRIORITY FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: US 09/106,926
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1320
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-779-144A-1

Alignment Scores:
Score: 5.05e-14 Length: 1320
Percent Identity: 203.50 Matches: 78
Percent Similarity: 37.99% Conservative: 58
Best Local Similarity: 21.79% Mismatches: 149
Query Match: 11.12% Indels: 73
DB: 10 Gaps: 12

US-09-857-613a-28 (1-350) x US-09-779-144A-1 (1-1320)
Qy 42 SerialaaserSerGluArgGlyLeuIleValleuGluGlnLysProLysLysAsp 61
Db 177 AACTCTGCCCCAAAGAAAGCCGTTGACAGAACTTGGAGATGATGAGAACGAT 236
Qy 62 LysLysLysLeuGlnLysGlyIleAlaGluPheTyrAspGluSerSerGlyLeuTyrp 81
Db 237 AAGATGCGCCAAAGAACGTCGCTT---GAGATATATATGAAGCCACATTCCTACTAT 293
Qy 82 AsnIle-----TrrpGlyAspHisMetHisGly---Phe 92
Db 294 AACGCTGTACAGATTTCTATGAATATGTTGGGCTTCCTTCCATTCAGAGATTT 353
Qy 93 TyrAspSerAspSer-----ThyValSerLeuSerAspHisArgAlaGlnIle 109
Db 354 TATAAAGGTGAGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 413
Qy 110 ArgMetIleGlnGluSerLeuArgPheAlaSerValSerGluGluArgSerLysTyrPro 129
Db 414 GCTGATTCACAAAGAGC----- 431
Qy 130 LysSerIleValAspValGlyCysGlyIleGlyLysSerArgTyrLeuAlaLysLys 149
Db 432 GATTAGTTCGACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 491
Qy 150 PheGlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAsnAla 169
Db 492 ACCGCTTGAACGCTACGCTTAAACATTAACGATTTACCAATTTGCCAAGCAAAATAT 551
Qy 170 LeuAlaAlaIleGlnGlyLeuAlaAspLysValSerPheGlnValAlaAlaLeuGln 189
Db 552 TACGCTTAAATAATTAACATTTGATGATGACCAAAATGCTTGTAAAGGCTGATTCATGAA 611
Qy 190 GlnProPheSerAspGlyGlnPheAspLeuValTyrPserMetGluSerGlyLeuHisMet 209
Db 612 ATGATTTTCGAGAAACACTTTCGACAAAGTTTATGCAATTTAGAGGCCATGTCACGCT 671
Qy 210 ProAspLysAlaLysPheValGlyLeuLeuAlaArgValAlaAlaProGlyAlaIleIle 229
Db 672 CCAAAATTAAGAGGCTATACAGCGCAATCTACAGATTTTGAACCGGCTGATCCTT 731
Qy 230 IleIleValIleThrPyrCysHisArgAspLeuGlyProAspGluGlnSerLeuHisProTyr 249
Db 732 GCTGTTTACGAATGGGATATGACTGATTAATATGACGAAACATTCCTGCAACAT----- 785

```

```

OY 250 GIUGLINSPLLEULYSYLIECYSPALATYTYR-----Leu 263
DB 786 -----AGAAAGATC-----GCTATGAATTCAGTACGTGATG 824
OY 264 PROALATPCYSSETHSERASPTRYVALYLSLEULGINSERLEUSERLEUGLNASP 283
DB 825 CCAAAAGATTCAGTACGTGCTAGCAAAAGCATGTGAAGAACTGGTTGCGAAGTC 884
OY 284 ILELYSSEGLNASP-----TRPSERARGPHEVALALA 294
DB 885 CTGGTTAGCGAAGACCTGGCGGACATGATGATAATCCCTGGATTATACCATTA 944
OY 295 PROPHETIPROALAVL-----LEARGSERALAPHE 305
DB 945 GGTGAGTGAAGTACGTTCAAAACCTAGCTAATTTGGCCACATTTTCAGAACTTCTAC 1004
OY 306 THRTPYSGLYLEUSER-----LEULUSER 315
DB 1005 TTGGGTAGACAAATTACTACAGCAATGTTACTGTATGAGAAAATTAGCTTACGCCCA 1064
OY 316 SERGLYGLNLSRTHRIEYLSGLYALALEU-----ALAMETPROLEUMETILEGLU 332
DB 1065 GAAGGTTCCAGAGCAATCTGCTGCTAGAAAATCTCGCTGGTTAGTGGCCGGT 1124
OY 333 GLTYTRYLSYLSASPLEULIELYSPHEALALELLETHRCYSARGLYSPROGLU 350
DB 1125 GGTAACTCCAAAGTTATTCACCTCAATGATGCTTTTGGCTGCTAGAGCCAGAA 1178

RESULT 8
US-09-779-144A-5
: Sequence 5, Application US/09779144A
: Patent No. US20020148006A1
: GENERAL INFORMATION:
: APPLICANT: Nes, David W.
: TITLE OF INVENTION: Transgenic Plants with Modified Sterol Compositions
: FILE REFERENCE: 11899, 0198, DVUS01 MOBT:198
: CURRENT APPLICATION NUMBER: US/09/779,144A
: CURRENT FILING DATE: 2001-02-08
: PRIOR APPLICATION NUMBER: US 09/106,926
: PRIOR FILING DATE: 1998-06-29
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: Patentin version 3.0
: SEQ ID NO 5
: LENGTH: 1320
: TYPE: DNA
: ORGANISM: Saccharomyces cerevisiae
: US-09-779-144A-5

Alignment Scores:
Pred. No.: 5 05e-14 Length: 1320
Score: 203.50 Matches: 78
Percent Similarity: 37.99% Conservative: 58
Best Local Similarity: 21.79% Mismatches: 149
Query Match: 11.12% Indels: 73
DB: 10 Caps: 12

US-09-857-613A-28 (1-350) x US-09-779-144A-5 (1-1320)
OY 42 SERALALASERGLUARGLYGLNLEVALLEUGLINSERPROLYSLYSPASP 61
DB 177 AACCTGCCCCAAAAGAACCGCTTCAGAACTTACGAAATTTGGATGATGATGATGAT 236
OY 62 LYSLSYLSYLSLEUGLINSGLYILEALAGLUHETRYASPLUSERSEGLYLEUTRPGLU 81
DB 237 AAAGATGCCAGAGAACGCTGCTT---GAGGATTATATATGAAGCCACACATTCCTACAT 293
OY 82 ASUNLE-----TRPGLYASPHISMETHISGLY---PHE 92
DB 294 AACGTCGTTACAGATTTTATGATATATGCTTGGCTTCCTTCCTTCATTTTCAGCATTT 353
OY 93 TYRASPSEASPSE-----THRVALSERLEUSERASPHISARGALALAGLNLLE 109

```

```

DB 354 TATTAAGTGACAGTTTGGCTGCTCGATGACAGACATGAACTTATTTAGCTTACAG 413
OY 110 ARGMETILEGLNLSERLEUARGPHEALASERVALSERGLUARGSERLYSTRPPO 129
DB 414 GCTGCTATTCAAAAGACC----- 431
OY 130 LYSSERIEVALASPVALCYLSGLYILEGLYGLYSERSEARGTRYLEUALAYSLS 149
DB 432 GATTGTCTCTGACCTGCTGTTGCTGTTGGGGCCCGACAGACAGATTCATGATTT 491
OY 150 PHEGLVALATHSERVALCYLILIEHTRLEUSERPROVALGIALAGLNRGALASNALA 169
DB 492 ACCGTTGTAGCTCATGCTGCTTAACAATAACGATTAACCAATTGCCAAGCAATAAT 551
OY 170 LEUALALALAGLNLLEUALAASPVYVALSERPHEGLNVALALASPALALEUGLN 189
DB 552 TACGCTAAAATAATACAAATTTGAGTACCAAAATGACCTTTGTAAGGGTATTCATGAAA 611
OY 190 GLNPROPHESERASPVLYGLNPHASPLEVALTRPSETMEGLUSERGLYUNHLSMET 209
DB 612 ATGGAATTCGAAAGAAAACACTTTCGACAAATTTATGCAATTTAGGCAATCTCACGCT 671
OY 210 PROASPLYSALALYSPHEVALGLYLULEUALAARGVALALALAPROGLYALALELLE 229
DB 672 CCAAAATTAGACGTATATACAGCAATCTACAGATTTTGAACCGGTGTACTCTT 721
OY 230 ILELEVALATHRTPCYSHISARGASPLEUGLYPROASPLUGLINSERLEUHSPTOTP 249
DB 732 GCTGTTTACGAATGGTAATGACTGATTAATATGACGAAAACATCTGTACAT 785
OY 250 GIUGLINSPLLEULYSYLIECYSPALATYTYR-----Leu 263
DB 786 -----AGAAAGATC-----GCTATGAATTCAGTACGTGATGATG 824
OY 264 PROALATPCYSSETHSERASPTRYVALYLSLEULGINSERLEUSERLEUGLNASP 283
DB 825 CCAAAAGATTCAGTACGTGCTAGCAAAAGCATGTGAAGAACTGGTTGCGAAGTC 884
OY 284 ILELYSSEGLNASP-----TRPSERARGPHEVALALA 294
DB 885 CTGGTTAGCGAAGACCTGGCGGACATGATGATAATCCCTGGATTATACCATTA 944
OY 295 PROPHETIPROALAVL-----LEARGSERALAPHE 305
DB 945 GGTGAGTGAAGTACGTTCAAAACCTTAGCTAATTTGGCCACATTTTCAGAACTTCTAC 1004
OY 306 THRTPYSGLYLEUSER-----LEULUSER 315
DB 1005 TTGGGTAGACAAATTACTACAGCAATGTTACTGTATGAGAAAATTAGCTTACGCCCA 1064
OY 316 SERGLYGLNLSRTHRIEYLSGLYALALEU-----ALAMETPROLEUMETILEGLU 332
DB 1065 GAAGGTTCCAGAGCAATCTGCTGCTAGAAAATCTCGCTGGTTAGTGGCCGGT 1124
OY 333 GLTYTRYLSYLSASPLEULIELYSPHEALALELLETHRCYSARGLYSPROGLU 350
DB 1125 GGTAACTCCAAAGTTATTCACCTCAATGATGCTTTTGGCTGCTAGAGCCAGAA 1178

RESULT 9
US-09-779-144A-3
: Sequence 3, Application US/09779144A
: Patent No. US20020148006A1
: GENERAL INFORMATION:
: APPLICANT: Nes, David W.
: TITLE OF INVENTION: Transgenic Plants with Modified Sterol Compositions
: FILE REFERENCE: 11899, 0198, DVUS01 MOBT:198
: CURRENT APPLICATION NUMBER: US/09/779,144A
: CURRENT FILING DATE: 2001-02-08
: PRIOR APPLICATION NUMBER: US 09/106,926
: PRIOR FILING DATE: 1998-06-29
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: Patentin version 3.0
: SEQ ID NO 3

```


OY	224	AlaproglylaIlellellelleValThrProCysHisArgAspGluProAspLys	243
Db	2520	AACCGAAGCATATTCCGCCTCACTACT-----ATCGGTTCGAAAAA	2564
OY	244	GlnSerLeuHis-----ProTrrpgLugInaspLeuLysLysIleCysAspAlaTyr	261
Db	2565	ACCAGATCTGAATCTTGATCTCCCTGG-----ATTAAATAAATAT	2600
OY	262	TyrlseuProAlatrrpcysserThrSeraSPtyrVallLysIleuGlnSerLeuSerLeu	281
Db	2601	ATTTTTCCGAACGGTTCCGCCCTCT-----GTACGCCAATTGGCTACGCCAGCAA	2654
OY	282	GlnaspIleLysSerGlnasptrrPserArghpheValAlaProPhetrrProAlaValIle	301
Db	2655	CCCCACTTTGTGATGCAAGACTGGCAATCACTTCGGTCTGATTAAC-----	2699
OY	302	ArgSerAlaPheThrTrpLysLysLeuSerSerLeuLeuSerSerGlyGlnLys	319
Db	2700	---GATACACTACGTTGATGGCGGTGTGAACAAGATTCCCTCGCCGCAATGGCAGAATAATCCG	2756
OY	320	---ThrlIeLysGlyAlaLeu-----AlaMetProLeuMet	330
Db	2757	GATACACTATAGTGAACGCTTTAAACAAGATGTTTACCTATTATTC	2799

```

RESULT 11
US-09-878-574-3629
: Sequence 3629, Application US/09878574
: Patent No. US20020110548A1
: GENERAL INFORMATION:
: APPLICANT: Byrum, Joseph R.
: APPLICANT: La Rosa, Thomas J.
: APPLICANT: Thompson, Michael D.
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
: TITLE OF INVENTION: Plants
: FILE REFERENCE: 38-21(15401)B
: CURRENT APPLICATION NUMBER: US/09/878,574
: CURRENT FILING DATE: 2001-12-21
: PRIOR APPLICATION NUMBER: 09/333,535
: PRIOR FILING DATE: 1999-06-14
: NUMBER OF SEQ ID NOS: 15775
: SEQ ID NO 3629
: LENGTH: 411
: TYPE: DNA
: ORGANISM: Glycine max
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (1)..(411)
: OTHER INFORMATION: unsure at all n locations
: OTHER INFORMATION: Clone ID: LIB3028-007-Q1-B1-H6
US-09-878-574-3629

```

Alignment Scores:	
Pred. No.:	2,11e-08
Score:	.148,50
Percent Similarity:	49.15%
Best Local Similarity:	33.05%
Query Match:	8.11%
DB:	10
	2
Length:	4111
Matches:	39
Conservative:	19
Mismatches:	50
Indels:	11
Gaps:	2

US-09-857-613A-28 (1-350) x US-09-878-574-3629 (1-411)

Oy 128 TTPProlysserIleValaspValaIGlyCYcsgIlyllecIglySerSerArqYrLeuAla 147

Db 3 TGGCCAAAAGTA- CTGAGATCTGGTGTGTGACAGAGAGAGGTGCATTTACATGCGT 60

Oy 148 LysLysPheGlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAla 167

Db 61 GAAAAATTTTGATGTTGAGGTTGTGGCAATTGACCTCTCC-----ATA 1020

Oy 168 Asn-----AlaLeuAlaAlaAlaGlnGlyLeuAlaAspLysValSerPheGln 1833

Db 103 AACGTCGATTTCTCTTGGCATTTGAACGCTCTATTGGACCTCAATCTCTGGTGAATTTTGA 1620

```

0y 184 ValAlaAspAlaLeuGlnGlnProPheSerTpsGlyGlnPheAspLeuValITPserMet 203
      |||||
Db 163 TGTGGCAATTCACATAAAAAACAATCCCTAGAAATCATTGTGTAATCTATATCCCGT 222
      |||||
0y 204 GluSerGlyGluHisMetProAspLysAlaLysPheValGlyGluLeuAlaArgValAla 223
      |||||
Db 223 GACGATTTGTCACATATAAAGATTAACCAACCACCTATTGAGATCATTTTACAATGGTTG 282
      |||||
0y 224 AlaProGlyAlaIleIleIleIleValThrProCysHisArgAspLeuGlyPro 241
      |||||
Db 283 AAGCTGGAGGTACACTTATTATTACGATATTACTGCAAAAGTGAAGCAAGTCCA 336
      |||||

```

```

RESULT 12
US-10-260-877-53
; Sequence 53, Application US/10260877
; Publication No. US20030021813A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Chovan, Linda E.
; APPLICANT: Hessler, Paul E.
; APPLICANT: Reich, Karl A.
; TITLE OF INVENTION: ESSENTIAL BACTERIA GENES AND GENOME
; TITLE OF INVENTION: SCANNING IN HAEMOPHILUS INFLUENZAE FOR THE IDENTIFICATION OF
; TITLE OF INVENTION: 'ESSENTIAL GENES'
; FILE REFERENCE: 6565.US.P1
; CURRENT APPLICATION NUMBER: US/10/260,877
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US/09/649,145
; PRIOR FILING DATE: 2000-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 753
; TYPE: DNA
; ORGANISM: H. influenzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(753)
; OTHER INFORMATION: HI-0095
US-10-260-877-53

```

Alignment Scores:	
Pred. No.:	4.02e-07
Score:	141.00
Percent Similarity:	44.68%
Best Local Similarity:	25.00%
Query Match:	7.70%
DB:	9
	7
	Gaps:
	5

US-09-857-613A-28 (1-350) x US-10-260-877-53 (1-753)

[illegible]

```

Db 418 TTATGCTT-----ACTCAGCATGTATGCTGGCGGAATGATCATCAACTAT--- 468
QY 248 PRTPLGLUGlnAspLeuLeuLysLysIleCysAspAlaTyrTyrLeuProAlaTrpCys 267
Db 469 -----CTAGAAATAATGCGCAAGCATTT----- 492
QY 268 SerThrSerAspTyrValLysLeuLeuGlnSerLeuSerLeuGlnAspIleLysSerGlu 287
Db 493 -----AAAGTACTGTACCGCATTAACGAAGAAT 522
QY 288 AspTrpSerAspPheValAlaProPheTrpProAlaValIleIleArgSerAlaPheTrp 307
Db 523 GGATGGAAAGCATATA-----TTCCAGAAAGTGTGTTAGAAATGTGTACTTTC 573
QY 308 LysGlyLeuSerSerLeuLeuSer 315
Db 574 TCTGTGAGATGATCATTTCTTC 597

RESULT 13
US-09-845-248-1
: Sequence 1, Application US/09845248
: Patent No. US20020042106A1
: GENERAL INFORMATION:
: APPLICANT: Nampoothiri, Madhavan
: TITLE OF INVENTION: Nucleotide Sequences Which Code for the CMA Gene
: FILE REFERENCE: 032301 WD 1161
: CURRENT APPLICATION NUMBER: US/09/845,248
: CURRENT FILING DATE: 2001-08-30
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 1801
: TYPE: DNA
: ORGANISM: Corynebacterium glutamicum
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (248)..(1600)
: OTHER INFORMATION:
US-09-845-248-1

Alignment Scores:
Pred. No.: 2,02e-06 Length: 1801
Score: 140.00 Matches: 83
Percent Similarity: 36.06% Conservative: 45
Best Local Similarity: 23.38% Mismatches: 120
Query Match: 7,658 Indels: 107
DB: 10 Gaps: 15

US-09-857-613a-28 (1-350) x US-09-845-248-1 (1-1801)
QY 53 LeuGlnGlnLysPro-----LysLysAspAsp 61
Db 674 ATGGAAACACCCCTCGCATGCGTAAAGCATGATCAACGGCGCTAGCATCCAGGACTCG 733
QY 62 LysLysLysLeuGlnLysGlyIleAlaGluPheTyrAspGlnSerGlyLeuTrpGlu 81
Db 734 AAATCCCGCACAAGAAAGCATGATGACTACACAGCTGGCGCATGACTTCTACTCC 793
QY 82 AsnIleTrpLysAspHisMetHis-----GlyPheTyrAspSerAspSerThrVal 99
Db 794 CTGTTTATGATGATTCACCATCTACCTGCGCGATTAATCAACGCCAGATCA--- 850
QY 100 SerLeuSerAspHisAlaGluAlaGlnIleArgMetIleGlnGlnSerLeuArgPheAla 119
Db 851 AGTTTGGAAACACCCCAAGAAACAAATACCGCTCATCTTTGAAACATGCGCTCG--- 907
QY 120 SerValSerGlnGluArgSerLysTrpProLysSerIleValAspValGlyCysGlyIle 139
Db 908 -----AAGAAGCGCATGCGCTCTCTACACGTGGCATGCGGTTGG 946
QY 140 GlyLysSerAspArgTyrLeuAlaLysLysPheGlyAlaThrSerValGlyIleThrLeu 159
Db 947 GGAGGATGTCGCTAC---GCCGCCAAACACGCTGTGAAGACCATGCGATTACGCTG 1003

```

```

QY 160 SerProValGlnAlaGlnArgAlaAsnAlaLeuAlaAlaGlnGlyLeuAlaAspLys 179
Db 1004 TCTGAACACCATATGAGTGGGCTCAAGCAGATCAACAGCCAGTTTGAAGACCTTC 1063
QY 180 ValSerPheGlnValAlaAspAlaLeuGlnGlnProPheSerAspGlyLysIlePheAspLeu 199
Db 1064 GCGGAATTCGCTTCATGATGATTAACCGCATGTCCA-----GAACTGATTCATGCG 1117
QY 200 ValTrpSerMetGlnSerGlyLysHisMet-----ProAsp----- 211
Db 1118 ATCTCAGCATATGGCATCTTGTAGACATCGGTGTAACAACATATCCGACACTTTCGA 1177
QY 212 -----LysAlaLysPheValGlyGluLeuAla----- 220
Db 1178 TTGCTCAGCAGCAAACTCAAAACAGCGCATGATCTCAACACAGCATCACTTACCA 1237
QY 221 -----ArgValAlaAlaProGlyAlaIleIleIleIleValThrTrpCysHisArgAsp 238
Db 1238 GACAAACCGCCCGCCGACGCGATGATTTATTT-----GATCGCTAC 1279
QY 239 LeuGlyProAsp----- 242
Db 1280 ATTTTCCCGAGCGTGAATCACTAGCTGCTTGGCACCCTGATCAAGCATGACGAGCAAC 1339
QY 243 -----GlnIleSerLeuHisPro 248
Db 1340 GTTTGGAGTGTGTCACGAGAAGAAACCTCCGCTTGTATTAACAGCACCCTGCACGGC 1399
QY 249 TrpGlnGlnAspLeuLeuLysLysIleCysAspAlaTyrTyrLeuProAla----- 265
Db 1400 TGCTGGGAAACCTCAAGAAATTTGGAGAGAGAGTGAACCTCCCGGTGAACCCACT 1459
QY 266 -----TrpCysSerThrSerAspTyrVal 273
Db 1460 GCACGACTCTTGGCTGATGATGCGATGCGAATGGGATTTGCCCAACATCTGC 1519
QY 274 LysLeuLeuGlnSerLeuSerLeuGlnAspIleLysSerGlnAspTrpSerArgPheVal 293
Db 1520 CACGTGACCAAGTACTGGTGTGAAA-----CTGATGACGAGGAGCACTCCGCGAGAA 1573
QY 294 AlaPro-----PheTrpProAlaValIleArg-SerAlaPheThrTrpLysGlyLeu 311
Db 1574 GTTCTGTAAGAAATGCTGCTGACTATCTAAAGAAACATGTTCTTTTAAAG----- 1625
QY 311 rSerLeuLeuSerSerGlyLysThrLysThrLysGlyAlaLeu 325
Db 1626 -----GTTCCCTAGGGGCGAGGATTATTTAGTAACTC 1661

RESULT 14
US-09-738-626-636
: Sequence 636, Application US/09738626
: Publication No. US20020197605A1
: GENERAL INFORMATION:
: APPLICANT: NAKAGAMA, SATOSHI
: APPLICANT: MIZOGUCHI, HIROSHI
: APPLICANT: ANDO, SEIKO
: APPLICANT: HAYASHI, MIKIRO
: APPLICANT: OCHIAI, KEIKO
: APPLICANT: YOKOI, HARUHIKO
: APPLICANT: TATEISHI, NAKO
: APPLICANT: SENOH, AKIHIRO
: APPLICANT: IKEDA, MASATO
: APPLICANT: OZAKI, AKIO
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
: FILE REFERENCE: 249-125
: CURRENT APPLICATION NUMBER: US/09/738,626
: CURRENT FILING DATE: 2000-12-18
: PRIOR APPLICATION NUMBER: JP 99/377484
: PRIOR FILING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: JP 00/159162
: PRIOR FILING DATE: 2000-04-07
: PRIOR APPLICATION NUMBER: JP 00/280988

```


GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame-Plus-p2n model

Run on: March 19, 2003, 06:53:46 : Search time 1289 Seconds
(without alignments) 4397.533 Million cell updates/sec

Title: US-09-857-613a-28
Perfect score: 1830
Sequence: 1 MATVRIPTISCIHHTFRS.....IEGKKDLKFAITCKRPE 350

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09857613/runat.12032003.100036.6027/app_query.fasta.1.519
-DB=EST -QFMT=fastap -SUFFIX=trst -MINMATCH=0.1 -LOOPTC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.ccl -LIST=45
-DOCALLIGN=200 -THR.SCORE=Pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=Pct -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09857613.ecgn.1.1456.ecgnat.12032003.100036.6027 -NCPU=6 -ICPU=3
-NO_XIPXY -NO_WMAP -LARGEOUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estinu:*
5: em_estlov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_iny:*
20: em_gss_pln:*
21: em_gss_vtc:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1136.5	62.1	1304	11 AY109799	AY109799 Zea mays
2	1029	56.2	599	14 BQ299297	BQ299297 sa045f11.
3	1024	56.0	669	10 BE331080	BE331080 so85a05.y
4	980	53.6	567	14 BQ273543	BQ273543 sa024e11.
5	975	53.3	570	14 BQ786302	BQ786302 saq66c02.
6	971.5	53.1	778	12 BG127606	BG127606 EST473168
7	964	52.7	554	13 BM527813	BM527813 sal166b10.
8	957	52.3	562	13 B1469324	B1469324 sal110d11.
9	946	51.7	556	13 B1469281	B1469281 sal109n01.
10	931	50.9	704	13 BQ288530	BQ288530 BQ288530
11	922	50.4	831	12 BG320098	BG320098 Zm03_01d1
12	899	49.1	664	13 BQ463750	BQ463750 BQ463750
13	884	48.3	515	14 BQ273302	BQ273302 sa023e12.
14	878	48.0	532	10 AM596072	AM596072 s197e09.y
15	869	47.5	628	9 A1487374	A1487374 EST245696
16	855	46.7	581	13 B1321569	B1321569 sal12h01.
17	844	46.1	509	12 BG511373	BG511373 sal17a10.
18	840	45.9	484	13 B1321583	B1321583 sal15a05.
19	836	45.7	708	14 BQ801716	BQ801716 WHE2817.G
20	824	45.0	503	12 BG790250	BG790250 sae67a12.
21	810	44.3	649	13 B1953869	B1953869 HVSME001
22	807	44.1	707	12 BQ416854	BQ416854 HVSME001
23	794	43.4	691	14 BQ116842	BQ116842 EST602418
24	788	43.1	540	10 BE331113	BE331113 so95906.y
25	777	42.5	517	12 BQ046436	BQ046436 saas4a02.y
26	775	42.3	587	10 BE340311	BE340311 EST344382
27	769	42.0	627	12 BG522228	BG522228 19-89 Ste
28	763	41.7	708	12 BQ116864	BQ116864 HVSME001
29	759	41.5	439	12 BE804492	BE804492 sr80d07.y
30	754	41.2	444	13 BM094790	BM094790 sa121d02.
31	734.5	40.1	653	10 AW774108	AW774108 EST33338
32	728	39.8	553	12 BF631792	BF631792 NF008D03D
33	720	39.3	587	10 BE591261	BE591261 WHE1655-1
34	710	38.6	661	13 BQ447900	BQ447900 BQ447900
35	707	38.6	733	12 BF051091	BF051091 EST436266
36	685.5	37.5	533	9 A1489112	A1489112 EST247451
37	670	36.6	517	10 BE607043	BE607043 WHE0915.H
38	665	36.3	484	10 AM617931	AM617931 EST314005
39	662.5	36.2	730	12 BG523936	BG523936 37-16 Ste
40	660	36.1	639	10 AV945401	AV945401 AV945401
41	658	36.0	491	9 A1485917	A1485917 EST244238
42	657	35.9	531	10 BE426891	BE426891 WHE033.C
43	656	35.8	452	10 AM099184	AM099184 sd35h03.y
44	650	35.5	623	12 BG447567	BG447567 NF004H01S
45	642	35.1	449	12 BG725508	BG725508 sae38c12.

ALIGNMENTS

RESULT 1
AY109799
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORANISM
Zea mays
Zea mays
Eukaryota: Viridiplantae: Streptophyta; Embryophyta: Tracheophyta;
Spermatophyta: Magnoliophyta: Liliopsida: Poales; Poaceae; PACC
Clade: Panicoideae; Andropogoneae; Zea.
I (bases 1 to 1304)
Hainey/C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.

TITLE Maize Mapping Project/Dupont Consensus Sequences for Design of Overgo Probes

JOURNAL Unpublished (2002)

REFERENCE 2 (bases 1 to 1304)

AUTHORS Coe, E.C.

TITLE Direct Submission

JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

FEATURES

Source

1. 1304

/organism="Zea mays"

/db.xref="MaizeDB:631709"

/db.xref="taxon:4577"

/clone="Cl364.1"

/clone_id="Maize Mapping Project/Dupont Consensus Library"

/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACS in conjunction with the Maize Mapping Project"

BASE COUNT 268 a 370 c 353 g 244 t 69 others

ORIGIN

Alignment Scores:

Pred. No.: 1,42e-106 Length: 1304

Score: 1136.50 Matches: 226

Percent Similarity: 71.87% Conservative: 32

Best Local Similarity: 62.95% Mismatches: 86

Query Match: 62.10% Indels: 15

DB: 11 Gaps: 4

US-09-857-613a-28 (1-350) x AY109799 (1-1304)

QY 4 ValValaGtIleProFhrIleSerCysIleHis-----IleHisThrPheArgSer 20

DB 28 ATATGCGGATCCCGGGCGGCAATGCTCACCGCGCGCTCCATTCG-----TCC 81

QY 21 GlnSerProArgThrPheAlaArgIleArgValGlyProArgSerTrpAlaProIleArg 40

DB 82 CACTCTCCAGAGCCCTCCAGCGCGCGGCGGCGAGCCACATACCGCCCTTCGCAC 141

QY 41 AlaSerAlaIleSerSerGluArgGlyIleValIleuGluGlnLysProLysLysAsp 60

DB 142 GTCCCGCGCCACTCCCGCGCTCCAGCGCGCGCTCGCTCGATCGCGATCGCTCG 201

QY 61 Asp-----LysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 72

DB 202 TCACAGCGCANN 261

QY 73 TyrAspGluSerSerGlyLeuTrpGluAsnIleTrpGlyAspHisMetHisIleGlyPhe 92

DB 262 TAGCAGAGTGTGCGGGGTGTGGAGAACATGTGGGGCGCACCATGACACAGCGCTTC 321

QY 93 TyrAspSerAspSerThrValSerLeuSerAspHisAlaAlaGlnIleArgMetIle 112

DB 322 TAGCACTCCAGGAGAGCGCCCTCCATGCGCATACCGCGCGCCAGATCGCATATC 381

QY 113 GlnGluSerLeuArgPheAlaSerVal-----SerGluGluArgSerLysTrpProLys 130

DB 382 GAGGAGCGCCCTTCCTCCCGCGGTCTCCAGCGCTCAGATGANNNNNNNNNNNNNNNN 441

QY 131 SerIleValAspValGlyCysGlyIleGlyGlySerSerArgTrpLeuAlaLysLysPhe 150

DB 442 NNNNNAGTCGATGTGAGATGTGCTGTGTAGTCAAGTACTTGGCGAAGAATATC 501

QY 151 GlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAsnAlaLeu 170

DB 502 GGAGCGCACTGCATCGCACCATGTTAGCCCTGTTCAGACCGGAGAGGAATATGCTTC 561

QY 171 AAtaAaIaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAlaLeuGlnGln 190

DB 562 GCTGCAGCGCAGGGGTTGTGCGATCAGTTACTTGCAGATTCGATCCTCGAGCAA 621

QY 191 ProPheSerAspGlyGlnPheAspLeuValTrpSerMetGluSerGlyGluHisMetPro 210

DB 622 CCGTTTCTCAGCGAGGATGTGATCTGTGTCTCCATGAGAGTGGCAGACATGCCG 681

QY 211 AspLysAlaLysPheValGlyGluLeuValAlaArgValAlaAlaProGlyAlaIleIle 230

DB 682 GACAGAGAAATTTTGTAGTACCTAGCACCGGCGGCTCCGTGAGGAGCAATATC 741

QY 231 IleValThrTrpCysHisArgAspLeuGluProAspGluGlnSerLeuHisProTrpGlu 250

DB 742 ATCGTACATGTCGTCATGAGAACTGCATGCAGAACTGCCTTAAGCCCATATA 801

QY 251 GlnAspLeuLeuLysLysIleCysAspAlaTrpTrpLeuProAlaTrpCysSerThr 270

DB 802 CTGACCTCTCCAGAGAGATATGACCGCGTACTCCCGAGTGTGCTCACTTCA 861

QY 271 AspTrpValLysLeuGlnSerLeuSerLeuGlnAspIleLysSerGluAspTrpSer 290

DB 862 GACTATGTGACATATGCCAAGTCACTGTCTCGAGATATCAAGACAGTCACTGTGCG 921

QY 291 ArgPheValAlaProPheTrpProAlaValIleArgSerAlaPheThrTrpLysGlyLeu 310

DB 922 GAGACGTGGCCCGCTTTTGGCCCGCGCTGATTAATATAGCGCTAACTGAGAGGCTTC 981

QY 311 SerSerLeuLeuSerSerGlyGlnLysThrIleLysGlyAlaLeuAlaMetProLeuMet 330

DB 982 ACCTCTGTGCTGAGCAGCGATGAGACGATGAGAGCGCGGATGATGATGATGATGATG 1041

QY 331 IleGluGlyTrpLysLysAspLeuIleLysPheAlaIleIleIleIleIleIleIleIle 349

DB 1042 ATCCAGGCTCAAGAGAGGCGCTCATCAATCAATCAATCAATCAATCAATCAATCA 1098

RESULT 2

B0299297 599 bp mRNA linear EST 16-MAY-2002

LOCUS sa045f11.y1 Gm-cl054 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-cl054-7197.5' similar to FR:09X1P9 09X1P9 GAMMA-TOCOPHEROL METHYLTRANSFERASE. [1], mRNA sequence.

ACCESSION B0299297

VERSION B0299297.1 GI:20814819

KEYWORDS EST.

SOURCE soybean.

ORGANISM Glycine max

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Rhanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, R., Steptoe, M., Rheising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schuck, R., Ritter, E., Kohn, S., Shin, T., Jackson, T., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.

PUBLIC Soybean EST Project

UNPUBLISHED (1999)

CONTACT: Shoemaker R./Public Soybean EST Project

PUBLIC Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccutresgen.com web site: www.resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 429.

Location/Qualifiers

1. 599

```

/organism="Glycine max"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-c1054-7197"
/clone_lib="Gm-c1054"
/tissue_type="leaf, 3 week old, greenhouse grown"
/lab_host="DH10B"
/Note="Vector: Bluescript II SK-, Site_1: EcoRI, Site_2:
XhoI. The Harosoy NIL was constructed and seed was
provided by Dr. J. Specht, University of Nebraska
(Shoemaker and Specht, 1995). The cDNA library was
constructed from mRNA isolated from leaf tissue at various
developmental stages of 3 week old greenhouse grown
plants. Complementary DNA was synthesized from mRNA using
a primer consisting of a poly(dT) sequence with a XhoI
restriction site and a 3' anchor. EcoRI adapters were
ligated to the blunt-ended cDNA fragments followed by
XhoI digestion. The cDNA fragments were directionally
cloned into the EcoRI-XhoI restriction site of the
pBluescript vector. The ligated cDNA fragments were
transformed into DH10B host cells (GibcoBRL). This
library was constructed in cooperation with Dr. Paul
Keim's laboratory at Northern Arizona University."
BASE COUNT      139 a      144 c      165 g      151 t
ORIGIN
Alignment Scores:
Pred. No.:      6,27e-96      Length:      599
Score:          1029.00      Matches:      199
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    56.23%      Indels:      0
DB:             14      Gaps:      0
US-09-857-613a-28 (1-350) x BQ299297 (1-599)
QY      6 ArgIleProThrIleSerCysIleHisIleHisThrPheArgSerGlnSerProArgThr 25
        |||||||
Db      3 AGGATCCCAACATCTCGATCGATCCACATCCACAGCTTCGTTCCATCCCTCGCAT 62
QY      26 PheAlaArgIleArgValGlyProArgSerTrpAlaProIleArgAlaSerAlaSer 45
        |||||||
Db      63 TTCGCCAACAATCCGGGTCGAGCCAGTCGTCGCTCATTCGGGCAATCCGACAGACC 122
QY      46 SerGluArgGlyGlnIleValLeuGlnGlnLysProLysAspAspLysLysLysLeu 65
        |||||||
Db      123 TCGGAGAGAGGGGAGATGATATTGAGACAGAGCCAGAGACAGATGACAAAGAGCGT 182
QY      66 GlnLysGlyIleAlaGlnPheArgPheArgPheArgPheArgPheArgPheArgPheArg 85
        |||||||
Db      183 CAGAAGGAGATCGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 242
QY      86 AspHisMetHisIleGlyPheArgPheArgPheArgPheArgPheArgPheArgPheArg 105
        |||||||
Db      243 GACGCAATGACACATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 302
QY      106 AlaAlaGlnIleArgMetIleGlnGlnSerLeuArgPheAlaSerValSerGlnGln 125
        |||||||
Db      303 GCTGTCAGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 362
QY      126 SerLysTrpProLysSerIleValAspValGlyCysGlyIleGlyLysSerArgTyr 145
        |||||||
Db      363 AGTAATATGCCCCAGAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 422
QY      146 LeuAlaLysLysPheGlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaGln 165
        |||||||
Db      423 CTGGCCAAAGAAATTTGGACCAACAGTGTAGCATCAGTCTGCTGCTGCTGCTGCTGCT 482
QY      166 ArgAlaAsnAlaLeuAlaAlaAlaGlnGlnLysLeuAlaAspLysValSerPheGlnAla 185
        |||||||
Db      483 AGAGCAAAATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 542
QY      186 AspAlaLeuGlnGlnProPheSerArgPheGlnGlnPheAspLeuValTrpSerMetGlu 204
        |||||||

```

```

Db      543 GACGCTACAGCAACATCTCTGACGCCAGTTTGATCTGTGTGCTCATGAG 599
RESULT 3
BE331080
LOCUS
DEFINITION
BE331080 669 bp mRNA linear EST 04-DEC-2001
so95a05.y1 Gm-c1041 glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-c1041-1065 5' similar to TR:092SK1 092SK1 GAMMA-TOCOPHEROL
METHYLTRANSFERASE, mRNA sequence.
ACCESSION
BE331080
VERSION
BE331080.1 GI:9204856
KEYWORDS
SOURCE
ORGANISM
soybean.
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustoids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 669)
REFERENCE
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna,
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,
Y., Pearson,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,
R., Rittler,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,
R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
TITLE
JOURNAL
COMMENT
Contact: Shoemaker R./Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccuteresgen.com
High quality sequence stop: 461.
FEATURES
source
location/Qualifiers
1..669
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1041-1065"
/clone_lib="Gm-c1041"
/tissue_type="Senescing leaves, mature plant, greenhouse
grown"
/lab_host="DH10B"
/Note="Vector: pTV73Pac (Pharmacia); Site_1: EcoRI;
Site_2: HindIII; This library was constructed from mRNA
isolated from senescing leaf tissue of mature greenhouse
grown plants of the cultivar Williams. Complementary DNA
was synthesized from mRNA using a 3' anchored poly(dT)
primer. EcoRI adapters were ligated to the blunt-ended
cDNA fragments followed by digestion with EcoRI and
HindIII. The cDNA fragments were directionally cloned
into the EcoRI-HindIII restriction site of the pTV73-Pac
vector. The ligated cDNA fragments were transformed into
DH10B host cells. This library was constructed by Dr.
Randy Shoemaker."
BASE COUNT      154 a      156 c      184 g      174 t      1 others
ORIGIN
Alignment Scores:
Pred. No.:      2.37e-95      Length:      669
Score:          1024.00      Matches:      200
Percent Similarity: 97.10%      Conservative: 1
Best Local Similarity: 96.62%      Mismatches: 6
Query Match:    55.96%      Indels:      0
DB:             10      Gaps:      0
US-09-857-613a-28 (1-350) x BE331080 (1-669)
QY      1 MetAlaThrValAlaArgIleProThrIleSerCysIleHisIleHisThrPheArgSer 20
        |||||||

```

Db 49 ATGCCACCGGTGGTGGAGATCCCAACATCTCATGCATCCACATCCACACAGTTCCGCTCC 108
 QY 21 GlnSerProArgThrPheAlaArgValGlyProArgSerThrPheAlaArg 40
 Db 109 CATGCCCTCCGACCTTCCCGCAAGATCCGGGTCGGCCAGGTCGTGGCTCTATTCCG 168
 QY 41 AlaSerAlaAlaSerSerGluArgGlyGluIleValLeuGluGlnIleProLysValAsp 60
 Db 169 GCATGGCGACGCGAGCTCGAGAGAGCGAGATAGTATTGGACAGAGAGCGAAGAGCAT 228
 QY 61 AspLysLysLysLeuGlnLysGlyIleAlaGluPheTyrAspLysSerGlyLeuTrp 80
 Db 229 GACAAGAAAGACCTGACAGAGCAATCCGACAGATTTCAGACGACGCTTCGCTTATCG 288
 QY 81 GluAlaAlaIleTrpGlyAspHisMetHisGlyPheTyrAspSerSerThrValSer 100
 Db 289 GAGAACATTTGGGGGCGACACCATGCGCATGGCTTTATGACCTGGATCCACTGTTTCG 348
 QY 101 LeuSerAspHisArgAlaAlaGlnIleArgMetIleGlnIleSerLeuArgPheAlaSer 120
 Db 349 CTTTCGATCATCGTGGCTCTCAGATCCGAATGCCAAGATCCAGAGCTCTTGGCTCTCT 408
 QY 121 ValSerGluGluArgSerLysTrpProLysSerIleValAspValGlyGlyIleGly 140
 Db 409 GTTCTGTAGAGACGTAAGTAAATGGCCCAAGAGTATAGTGTGGTGGTGCATANGT 468
 QY 141 GlySerSerArgTyrIleuAlaLysLysPheGlyAlaThrSerValGlyIleThrLeuSer 160
 Db 469 GGCACCTCTAGTACTGCGCCAGAAATTTGGAGCACACAGTGTATGCATCTGAGT 528
 QY 161 ProValGlnAlaGlnArgAlaAsnAlaLeuAlaAlaGlnIleGlnIleuAlaAspLysVal 180
 Db 529 CCGTTCAAGCTCATAGACACATGCTTGTGCTGCTGCTCAAGAGATTGGCTGATATGTT 588
 QY 181 SerPheGlnValAlaAlaAspAlaLeuGlnGlnIleProPheSerAspGlyGlnPheAspLeuVal 200
 Db 589 TCCCTTCATGTGCTGACGCTCTACAGCAACCATCTCTGACGGCGAGTTGATCTGTG 648
 QY 201 TrpSerMetGluSerGlyVal 207
 Db 649 TCGTTCATGACAGATGTCGAGAG 669

RESULT 4
 LOCUS BQ273543 567 bp mRNA linear EST 07-MAY-2002
 DEFINITION sac24ell.y1 Gm-cl054 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-cl054-5206.5, similar to TR:09XTP9 09XTP9 GAMMA-TOCOPHEROL METHYLTRANSFERASE. [1] ; mRNA sequence.

ACCESSION BQ273543
 VERSION BQ273543.1 GI:20498613
 KEYWORDS EST.
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Virdiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 567)
 AUTHORS Shoemaker R., Kelm P., Vodkin L., Eipelting J., Corryell V., Khanna A., Bolla B., Marra M., Hillier L., Kucaba T., Martin J., Beck C., Wylie T., Underwood K., Steptoe M., Theising B., Allen M., Bowers Y., Person B., Swaller T., Gibbons M., Pape D., Harvey N., Schurk R., Ritter E., Kohn S., Shin T., Jackson Y., Cardenas M., McCann R., Waterston R. and Wilson R.
 TITLE Public Soybean EST Project
 JOURNAL Unpublished (1999)
 COMMENT Contact: Shoemaker R./Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: c@resgen.com web site: www.resgen.com
 Seq primer: -40RP from Gibco
 High quality sequence stop: 420.
 Location/Qualifiers
 1. 567
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="SOYBEAN CLONE ID: Gm-cl054-5206"
 /clone.lib="Gm-cl054"
 /tissue.type="Leaf, 3 week old, greenhouse grown"
 /lab.host="DH10B"
 /note="vector: pBluescript II SK+; Site_1: EcoRI; Site_2: XhoI; The Harosoy NIL was constructed and seed was provided by Dr. J. Specht, University of Nebraska (Shoemaker and Specht, 1995). The cDNA library was constructed from mRNA isolated from leaf tissue at various developmental stages of 3 week old greenhouse grown plants. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site and a 3' anchor. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in cooperation with Dr. Paul Kelm's laboratory at Northern Arizona University."

BASE COUNT 133 a 128 c 144 g 162 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 6,56e-91 Length: 567
 Score: 980.00 Matches: 187
 Percent Similarity: 99.47% Conservative: 0
 Best Local Similarity: 99.47% Mismatches: 1
 Query Match: 53.55% Indels: 0
 DB: 14 Gaps: 0

US-09-857-613a-28 (1-350) x BQ273543 (1-567)

QY 89 HISHISGLYPheTyrAspSerAspSerThrValSerLeuSerAspHisArgAlaGln 108
 Db 3 CACCATGGCTTTATGACCTGGAGATTCACCTGTTCCCTTCGGATATGCTGCTCAG 62
 QY 109 IleArgMetIleGlnGluSerLeuArgPheAlaSerValSerGluGluArgSerLysTrp 128
 Db 63 ATCCGATGATCCAGAGCTCTTCCCTTGGCTCTGCTTTCAGAGAGCTAGTAATG 122
 QY 129 ProLysSerIleValAspValGlyGlyIleGlyIleGlySerSerArgTyrLeuAlaLys 148
 Db 123 CCCAAGAGTATGTTGATGTTGGGTGCGCAVAGGCGACGCTCTGATVACCTGCCAAG 182
 QY 149 LysPheGlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaGln 168
 Db 183 AATTTTGGAGCAACCGTGTAGGCTACTGTAGCTCTGTTCAAGCTCAAGAGCAAT 242
 QY 169 AlaLeuAlaAlaGlnIleuAlaAspLysValSerPheGlnValAlaAspAlaLeu 188
 Db 243 GCTCTTGGCTGCTGCTCAAGAGATTGGCTGATAGGTTCTTTCAGGTTCTGACGCTCA 302
 QY 189 GlnGlnProPheSerAspGlyGlnPheAspLeuValTrpSerMetGluSerGlyGluHis 208
 Db 303 CAGCAACCATTTCTGACGCGCAGTTGATGCTGTGCTCATGAGAGAGAGCAT 362
 QY 209 MetProAspLysAlaLysPheValGlyGluLeuAlaArgValAlaAlaProGlyAlaIle 228
 Db 363 ATGCCGACAAAGCTTAGTTTGTGAGAGTTAGCTGGGTGACACACAGGTGCTCACT 422
 QY 229 IleIleIleValThrTrpCysHisArgAspLeuGlyProAspGluGlnSerLeuHisPro 248

ORIGIN

Alignment Scores:	
Pred. No.:	2.84e-89
Score:	964.00
Percent Similarity:	99.46%
Best local Similarity:	99.46%
Query Match:	52.68%
DB:	13
Length:	554
Matches:	183
Conservative:	0
Mismatches:	1
Indels:	0
Gaps:	0

US-09-857-613A-28 (1-350) x BM527813 (1-554)

Oy		68	GlytlaelaaglupehyrAspGluSerSercglyLeutrPgluaSnllleTroplysAphts	87
Dd		2	GGATTCGCAGAGTTTACGACGAGCTCTTGCGTTATGCGAACAATTGGGCCACCAC	61
Oy		88	MethlshISgIypheryrAspSeraSpserThValSerLeuSerAspHisArglaala	107
Dd		62	ATGCACCATGTGCTTTATGACTCGGATTCACCACTGTTTCGCTTCGATCATTGCTGCT	121
Oy		108	GlnlleatgmellieglnglunSerleuarqPhelaaservaIselSerGlunglarSericlys	127
Dd		122	CAGATCCGAAGAATCCAAGAGCTCTTCGCTTTCGCTTCGTTCTGAGAGCGTAGTA	181
Oy		128	TrrProlySserllevalAspValGlcySGlylieglLyglYSerSeraTrtyrLeuala	147
Dd		182	TGCCCCAACAGATAATGATTGATGTGGGTGTGGCAATAGCTGGACGCTTATGATACCTGGCC	241
Oy		148	LysLysPhecgIyalalArthrSerValGlYlIerThrLeuSerProvalGlInlaaglInargala	167
Dd		242	AAGAAATTTGGAGCAACCAAGTGTAGCAVCACCTGTGACGCTTCACAGCTCAAAGAGCA	301
Oy		168	AsnAlaLeualAlaAlaAlaGlnglyeunaLaaspLysValSerPheglInValaLaaspala	187
Dd		302	AATGCTCTTGGCTGCCTCAAGAGATTGGCTGATTAAGGTTTTCTTTCAGAGTTTCTGCAGCT	361
Oy		188	LeuglnglnnPropheSerAspPolylcInPheaSplauValTrrSerMetcluSerGlygu	207
Dd		362	CTACAGCAACCATTTCTCTGACAGGCCAGATTGTATCTGTGGTGTGCTCCATGAGAGTGGACAG	421
Oy		208	HismetProaspLysAlaLysPheValGlYgluLeualaArqValaAlaLabProglYala	227
Dd		422	CATATGGCTGACAAAGCTAAGTTTGTGGAGAGTTAGCTCGGGTAGCACACAGAGTGCC	481
Oy		228	IlellellellelvalThTrrPCySHisArgaspLeugLYProaspGluInSerleuHls	247
Dd		482	ACTAATAAATAATVAGTAAACATGATGTCACAGAGGACTCTTGCCCTGACAGAACAACTTACAT	541
Oy		248	ProttrpgLuIn 251	
Dd		542	CCATGGGAGCA 553	
<hr/>				
RESULT 8				
LOCUS	B1469324	562 bp	mRNA	linear EST 30-NOV-2001
DEFINITION	sali00411.y1 Gm-cl053 glycine max cDNA clone GENOME SYSTEMS CLONE			
	ID: Gm-cl053-2854.5, similar to tr:O9XIP9 O9XIP9 GAMMA-TOCOPIEROL			
	MEPHYLTRANSFERASE. [1] ; mRNA sequence.			
ACCESSION	B1469324			
VERSION	B1469324.1			
KEYWORDS	GI:15285433			
SOURCE	EST.			
ORGANISM	soybean.			
	glycine max			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
	Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae			
	Glycine.			
REFERENCE	1 (bases 1 to 562)			
AUTHORS	Shoenaker,R., Keim,P., Vodkin,L., Erpellding,J., Coryell,V., Khanna,			
	A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.C.,			
	Wylie,T., Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers,			
	J., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schuck			
	,R., Rltter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McClann			

TITLE
JOURNAL
COMMENT

P., Waterston, R. and Wilson, R.
 Public Soybean EST project
 Unpublished (1999)
 Contact: Shoemaker R./Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewaterston.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 for further information
 call: (800)-533-4363 or contact via email: ccu@resgen.com
 High quality sequence stop: 421.

FEATURES

SOURCE

/organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl053-2654"
 /clone_11b="Gm-cl053"
 /tissue_type="Whole seedling, 3 week old, greenhouse
 grown"
 /lab_host="DH10B"
 /note="vector: plusscript II SK+ Site_1: EcoRI; Site_2
 XhoI; The Harosoy NIL was constructed and seeded was
 provided by Dr. J. Specht, University of Nebraska
 (Shoemaker and Specht, 1995). The cDNA library was
 constructed from mRNA isolated from whole seedlings of 3
 week old greenhouse grown plants. Complementary DNA was
 synthesized from mRNA using a primer consisting of a 3'
 poly(dT) sequence with a XhoI restriction site and a 3'
 anchor. EcoRI adapters were ligated to the blunt-ended
 cDNA fragments followed by XhoI digestion. The cDNA
 fragments were directionally cloned into the EcoRI-XhoI
 restriction site of the plusscript vector. The ligated
 cDNA fragments were transformed into DH10B host cells
 (GibcoBRL). This library was constructed in cooperation
 with Dr. Paul Keim's laboratory at Northern Arizona

BASE COUNT
ORIGIN

130 a	136 c	153 g	143 t
-------	-------	-------	-------

Alignment Scores:	
Pred. No.:	1,52e-88
Score:	957.00
Percent Similarity:	99.47%
Best Local Similarity:	99.47%
Query Match:	52.30%
DB:	13
Length:	562
Matches:	186
Conservative:	0
Mismatches:	1
Indels:	0
Gaps:	0

US-09-857-613A-28 (1-350) x B1469324 (1-562)

Qy	11	SerGysIleHisIleHisThrPhaArgSerGlnSerProaGlyThrPheAlaArgIleArg	30
Db	2	TcATCATCCATCCATCCACACAGCTTCGTTCCCAATCCCCCGACATTTCCGCGAATCCGG	61
Qy	31	ValGlyProArgSerTrpPalaProIleArgAlaSerAlaAlaSerSerGluArgGlyGlu	50
Db	62	GTCGACACCCAGCTCGTGGCGCTCTATTATCGGCATTCGCGACCGACTCGGAAAGAGGGGAG	121
Qy	51	IleValIleuGluGlnIlyProIlyIysAspAspIlyIyIyIyIeGluGlnIyAla	70
Db	122	ATAGATTGGACCAACACCCAGAGAGATATACAGAGAAAGCTGCGACAAAGGAAATGCCA	188
Qy	71	GluPheTrpAspGluSerSerGlyLeuTrpGluAsnIleTrpGlyAspHisMetHisHis	90
Db	182	GAGTTTACGAGAGACGTCCTTGCGTTATGCGAGAAATTGGGGCGACCAATCCACCAT	244
Qy	91	GlyPheTrpAspSerAspSerThrValSerIleuSerAspHisArgAlaAlaGlnIleArg	110
Db	242	GCGTTTATGACTCGGATTCACACTTTTCCTTCGATATCCTGTCGCTCGACAAATCCGA	300
Qy	111	MetIleGlnGluSerLeuArgPheAlaSerValSerGluGluArgSerIlyTrpProIys	130

```

|||||
Db 302 ATGATCCAAAGAGTCTCTCCCTTGTCTGAGGAGCGTAGTAATGAGCCCAAG 361
Oy 131 SerIleValAspValGlyCysGlyIleGlyGlySerSerArgTyrIleuAlaLysLysPhe 150
Db 362 AGATAGTCTGATGTTGGTGTGGATAGTGGACGCTTCAATCATCTGGCCCAAGAAATT 421
Oy 151 GlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAsnAlaLeu 170
Db 422 GGAGCAACCAAGCTAGACATCAGCTGAGTCTTCAAGCTCAAGACCAATGCTCTT 481
Oy 171 AlaAlaIaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAlaLeuGln 190
Db 482 GCTGCTGCTCAAGGATGGCTGATAGATTCTTCAGGTTGCTGACGCTCTACAGCAA 541
Oy 191 ProPheSerAspGlyGlnPhe 197
Db 542 CCATTCTCTGACGGCAGTTT 562

RESULT 9
BI469281
LOCUS 556 bp mRNA linear EST 30-NOV-2001
DEFINITION sai09h01.y1 Gm-cl053 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl053-3025 5' similar to TR:09x1p9 09x1p9 GAMMA-TOCOPHEROL
METHYLTRANSFERASE. [1] ; mRNA sequence.
BI469281
ACCESSION BI469281.1 GI:15285390
VERSION EST.
KEYWORDS soybean.
SOURCE Glycine max
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 556)
Shoemaker, R., Kelm, P., Vodka, L., Erpelting, J., Coryell, V., Khanna
, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wyle, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers
, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: cures@resgen.com
High quality sequence stop: 422.
Location/Qualifiers
1. 556
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl053-3025"
/clone_id="Gm-cl053"
/tissue_type="Whole seedling, 3 week old, greenhouse
grown"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+. Site.1: EcoRI; Site.2:
XhoI; The Harosoy NIL was constructed and seed was
provided by Dr. J. Specht, University of Nebraska
(Shoemaker and Specht, 1995). The cDNA library was
constructed from mRNA isolated from whole seedlings of 3
week old greenhouse grown plants. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with a XhoI restriction site and a 3'
anchor. EcoRI adapters were ligated to the blunt-ended
cDNA fragments followed by XhoI digestion. The cDNA

```

```

Fragments were directionally cloned into the EcoRI-XhoI
restriction site of the pBluescript vector. The ligated
cDNA fragments were transformed into DH10B host cells
(GibcoBRL). This library was constructed in cooperation
with Dr. Paul Keim's laboratory at Northern Arizona
University."
BASE COUNT 129 a 135 c 152 g 140 t
ORIGIN
Alignment Scores:
Pred. No.: 2,04e-87 Length: 556
Score: 946.00 Matches: 184
Percent Similarity: 99.46% Conservative: 0
Best local Similarity: 99.46% Mismatches: 1
Query Match: 51.69% Indels: 0
DB: 13 Gaps: 0
US-09-857-613a-28 (1-350) x BI469281 (1-556)
Oy 11 SerCysIleHisIleHisThrPheArgSerGlnSerProArgThrPheAlaArgIleArg 30
Db 2 TCATGCATCCACATCCACACAGCTCCGCTTCCCAATCCCTCGACCTTGGCCGATCCGG 61
Oy 31 ValGlyProArgSerThrAlaProIleArgAlaSerAlaAlaSerSerGluArgGlyGlu 50
Db 62 GTCGACCCACAGCTCGTGGCTCTTATTCGGCATCGGACGACGCTCGGAGAGAGGGGAG 121
Oy 51 IleValLeuGlnGlnLysProLysAspLysLysLysLysLysLysLysLysLysLysLys 70
Db 122 ATAGATTGGACAGCAAGCCAGAAAGATGACAGAAAGAACTCCAGAAAGGAAATCCGA 181
Oy 71 GluPheTyrAspGlnSerSerGlyLeuTrpGluAsnIleTrpGlyAspHisMetHisHis 90
Db 182 GAGTTTACGACAGACTCTTCTGCTTATGAGGAGACATTTGGGGGAGACATGACCAT 241
Oy 91 GlyPheTyrAspSerAspSerThrValSerLeuSerAspHisArgAlaAlaGlnIleArg 110
Db 242 GCTTTTATGACTCCGATTCCTCCTGCTTGGATCATCGCTGCTCAGATCCGA 301
Oy 111 MetIleGlnLysLeuArgPheAlaSerValSerGlnLysArgSerLysTrpProLys 130
Db 302 ATGATCCAAAGAGTCTCTTCTGCTTCTGCTTCTGAGGACGATGATTAATGAGCCCAAG 361
Oy 131 SerIleValAspValGlyCysGlyIleGlyGlySerSerArgTyrIleuAlaLysLysPhe 150
Db 362 ACTATAGTGTGATGTTGGGTGGATAGCTGACAGCTGAGTACCTGCTCCAAAGAAATT 421
Oy 151 GlyAlaThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAsnAlaLeu 170
Db 422 GGAGCAACCAAGCTAGACATCAGCTGAGTCTTCAAGCTCAAGACCAATGCTCTT 481
Oy 171 AlaAlaIaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAlaLeuGln 190
Db 482 GCTGCTGCTCAAGGATGGCTGATAGATTCTTCAGGTTGCTGACGCTCTACAGCAA 541
Oy 191 ProPheSerAspGly 195
Db 542 CCATTCTCTGACGGC 556

RESULT 10
LOCUS 704 bp mRNA linear EST 09-APR-2002
DEFINITION BJ288530 Y. Ogihara unpublished cDNA library, Wh-SL Triticum
aestivum cDNA clone whs114p22 5', mRNA sequence.
ACCESSION BJ288530
VERSION BJ288530.1 GI:20106970
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae
; Triticaceae; Triticum.
REFERENCE 1 (bases 1 to 704)

```

AUTHORS Ogihara, Y. and Murai, K.
TITLE Expressed genes in *Triticum aestivum*
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i

FEATURES	Location/Qualifiers
source	1. : 704

Alignment Scores:	
Pred. No.:	9,74e-86
Score:	931.00
Percent Similarity:	85.46%
Best Local Similarity:	76.21%
Query Match:	50.87%
DB:	13
Length:	704
Matches:	173
Conservative:	21
Mismatches:	33
Indels:	0
Gaps:	0

US-09-857-613A-28 (1-350) x BJ288530 (1-704)

Oy	83	IIleptpGLysAspHisMetHisHisGlyPheThrAspSerAspSerThrValSerLeuSer	102
Db	22	ATCGGGGGCGGCAACATGCACACAGGGTTTACGACTCCGGCGAGCCCGCTCCATGTCC	81
Oy	103	AspHisArgAlaIaIaGlnIIleArgMetIIleGlnGluSerLeuArgPheAlaSerValSer	122
Db	82	GACCAACGGCGCGCCGAGATCCGGATCATGTAGAGAGGGCCCTGGCCCTTCCGCGCGCTCC	141
Oy	123	GluGluArgSerIyStrpProLysSerIIleValAspValGlyCysGlyIIleGlyGlySer	142
Db	142	GACCATCCGACAAACCAACCCGAAACCATTTGTGATGTCGGATCCGAAATCGGTGTAC	201
Oy	143	SerArgIyIleuAlaLysLysPheGlyAlaIaIaThrSerValGlyIIleThrLeuSerProVal	162
Db	202	TCAAGATACCTGGGCAACAAATATGGAGCAACATGCTCTCGGATTCACATTGAGCCACAGTg	261
Oy	163	GlnAlaGlnArgAlaAsnAlaLeuAlaIaIaAlaGlnGlyLeuAlaAspLysValSerPhe	182
Db	262	CAAGCCGAGAGAGGAATGCCCTTCGACAGCACACAGGGTTGTCCGACAAGGCTTCTTTC	321
Oy	183	GlnValAlaAspAlaLeuGlnGlnProPheSerAspGlyGlnPheAspLeuValITrpSer	202
Db	322	CAAGTGTGTATGCTCTGGAGCAACCATTTCCATGTGCGCAGTTTATCTTTCCTGTGCT	381
Oy	203	MetGluSerGlyGluHisMetProAspLysAlaLysPheValGlyIleuAlaIaIaArgVal	222
Db	382	ATGGAGACTGGTGGACCATCTCCCAACAAACAGAAAGTTTGAACGAGGCTGGCACGGGTC	441

[illegible]

RESULT	11			
LOCUS	BG320098/c			
DEFINITION	BG320098	831 bp	mRNA	linear
ACCESSION	Zm03.01d11_A	Zm03.AAFC_ECCRC_cold-stressed_maize-seedlings	Zea mays	EST 27-FEB-2001
VERSION	BG320098	cdna clone Zm03.01d11, mRNA sequence.		
KEYWORDS	BG320098.1	GI:13149776		
SOURCE	EST.			
ORGANISM	Zea mays			

RESEARCH

Source

	a	c	g	t	others
BASE COUNT	166	237	213	214	1
ORIGIN					

Alignment Scores:	
Pred. No.:	1.02e-84
Score:	922.00
Percent Similarity:	84.19%
Best Local Similarity:	74.36%
Query Match:	50.38%
DB:	12
Length:	831
Matches:	173
Conservative:	23
Mismatches:	37
Indels:	0
Gaps:	0

US-09-857-613a-28 (1-350) x BG320098 (1-831)

QY 116 LeuArgPheAlaSerValSerGluArgSerIleTrpProlySerIleValAlaPyl 135
 DB 827 CTTCGCCGCTGCCAGCTCCAGATATCCAGAGACACCAAAAATAGTCATGTC 768
 QY 136 GlycSerGlyIleGlyGlySerSerArgTyrlLeuAlaLysLysPheGlyAlaThrSerVal 155
 DB 767 GGATTCGCATTTGGTGGTACGCTCAAGTACTTGGCGAAGAAATACGAGCCAGTCAC 708
 QY 156 GlyIleThrLeuSerProValGlnAlaGlnArgAlaAsnAlaLeuAlaAlaGlnGly 175
 DB 707 GGGATCAGCTTACGCCCTCTCAAGCCGAGAGAGAAATGCTCTCCGTCGACGCCAGGG 648
 QY 176 LeuAlaAspLysValSerPheGlnValAlaAspAlaLeuGlnGlnProPheSerSply 195
 DB 647 TTTCGCGATCAGGTACCTCTCAAGTCTGATGCTGAGCAACCTTTCCTGACGG 588
 QY 196 GlnPheAspLeuValTrpSerMetGluSerGlyGlnHisMetProAspLysAlaLysPhe 215
 DB 587 CAGTTCGATCTGTGTGTGTCCATGAGAGTGGCCGACGACATGCCGACAGAGAAAGTTT 528
 QY 216 ValGlyGluLeuAlaArgValAlaAlaProGlyAlaIleIleIleValThrTrpCys 235
 DB 527 GTTAGTGAGCTAGACGCGCTGCGCTCTGAGAGGACAAATATCATCTGACATGTGC 468
 QY 236 HisArgAspLeuGlyProAspGluGlnSerLeuHisProTrpGluGlnAspLeuLys 255
 DB 467 CATAGAACTCGATGATCCGAAACCTGCTAAAGCCCGATGAACTAGCCTCTCTGAGG 408
 QY 256 LysIleCysAspAlaTyrlTyrlLeuProAlaTrpCysSerThrSerAspTyrlValLysLeu 275
 DB 407 AGGATATCGACGCGCTACTACCTCCGACGCTGCTCACCCTGACATATGTGAACATT 348
 QY 276 LeuGlnSerLeuSerLeuGlnAspIleLysSerGluAspTrpSerArgPheValAlaPro 295
 DB 347 GCCAAGTCACTGCTCTCTGAGATATCAAGACAGCTGCTCGAGAACCTGCGCCCG 288
 QY 296 PheTrpProAlaValIleArgSerAlaPheThrTrpLysGlyLeuSerSerLeuLeuSer 315
 DB 287 TTTTGCCCGCCGCTATTAATACAGCCGCTAACATGGAAGGGCTTACCTCTGCTGACG 228
 QY 316 SerGlyGlnLysThrIleLysGlyAlaLeuAlaMetProLeuMetIleGluGlyTyrlLys 335
 DB 227 ACCGATGGAAGACGATCAGAGCCCGATGGATGCGCTAATGATCCAGGCTACAAAG 168
 QY 336 LysAspLeuIleLysPheAlaIleIleThrCysArgLysPro 349
 DB 167 AAGGGCTCATCAATTCACCATCATCTCTCCCAAGCCT 126
 RESULT 12
 LOCUS BJA63750 664 bp mRNA linear EST 23-MAY-2002
 DEFINITION BJA63750 K. Sato unpublished cDNA library, cv. Haruna Nijo
 germination shoots Hordeum vulgare subsp. vulgare cDNA clone
 baas30f10 5', mRNA sequence.
 ACCESSION BJA63750
 VERSION BJA63750.1 GI:21142257
 KEYWORDS EST.
 SOURCE Hordeum vulgare subsp. vulgare.
 ORGANISM Hordeum vulgare subsp. vulgare.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
 ; Triticeae; Hordeum.
 REFERENCE 1 (bases 1 to 664)
 AUTHORS Sato, K., Saitoh, D. and Takeda, K.
 TITLE Barley EST sequencing project in NIG and Okayama Univ
 JOURNAL Unpublished (2002)
 COMMENT Contact: Tadasu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856

FEATURES
 source
 1. 664
 /organism="Hordeum vulgare subsp. vulgare"
 /cultivar="Haruna Nijo"
 /db_xref="taxon:112509"
 /clone="baas30f10"
 /clone_lib="K. Sato unpublished cDNA library, cv. Haruna
 Nijo germination shoots"
 /tissue_type="shoots"
 /dev_stage="germination"

BASE COUNT 158 a 177 c 186 g 143 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.8e-82 Length: 664
 Score: 899.00 Matches: 168
 Percent Similarity: 86.24% Conservative: 20
 Best local Similarity: 77.06% Mismatches: 30
 Query Match: 49.13% Indels: 0
 DB: 13 Gaps: 0

US-09-857-613a-28 (1-350) x BJA63750 (1-664)

QY 73 TyrAspGluSerSerGlyLeuTrpGluAsnIleTrpGlyAspHisMetHisGlyPhe 92
 DB 4 TACAGCAGATGTCGCCGCTGTGGAGACATCTGGGCGACACATCCACCGCTTC 63
 QY 93 TyrAspSerAspSerThrValSerLeuSerAspHisArgAlaAlaGlnIleArgMetIle 112
 DB 64 TACGACTCTGGGAGACCGCCCTCCATGCTCCGACACCCCGGCCGACATGCCATGATC 123
 QY 113 GlnGluSerLeuArgPheAlaSerValSerGluGlnArgSerLysTrpProlySerIle 132
 DB 124 GAGGAGGCCCTGCTTCCGCGCCCTCCCGACGACCGGACAAACAAACAAATTT 183
 QY 133 ValAspValGlyCysGlyIleGlyGlySerSerArgTyrlLeuAlaLysLysPheGlyAla 152
 DB 184 GTTATGTTGGATGGGAAATCGCGGCTAGCTCAAGTACCTGGCAACAAATATGAGACA 243
 QY 153 ThrSerValGlyIleThrLeuSerProValGlnAlaGlnArgAlaAsnAlaLeuAla 172
 DB 244 CAGTCTCTGGGATCACAATTGAGCCAGCTGCAAGCCGAGAGGAAATGCCCTACCGCG 303
 QY 173 AlaGlnGlyLeuAlaAspLysValSerPheGlnValAlaAspAlaLeuGlnProPhe 192
 DB 304 GCACAGGGGTGGCAGACAGGCTTTTCCAGTTGCTGATGCTGTGAGCAACCATTT 363
 QY 193 SerAspGlyGlnPheAspLeuValTrpSerMetGluSerGlyGlnHisMetProAspLys 212
 DB 364 CCTGATGGGCACTTGTGATCTGTGTGTATGAGAGTGGTGAACATGCCGACAAADA 423
 QY 213 AlaLysPheValGlyGluLeuAlaArgValAlaAlaProGlyAlaIleIleIleVal 232
 DB 424 CAGAGCTTTGTAGTGAGCTGCACGCGCTCCACGCTCCAGGAGCAACATATCATCTGTG 483
 QY 233 ThrTrpCysHisArgAspLeuGlyProAspGluGlnSerLeuHisProTrpGluGlnAsp 252
 DB 484 ACCTGCTCCATAGGAGATCCCTGCTATCTGAGACTGATGAACCTGAGCAAGCTGAT 543
 QY 253 LeuLeuLysLysIleCysAspAlaTyrlTyrlLeuProAlaTrpCysSerThrSerSply 272
 DB 544 CTTTGAAGAAAGATTTGGTATCATATTAATCCCGGATGTGTCTCCCGCGGATAT 603
 QY 273 ValLysLeuLeuGlnSerLeuSerLeuGlnAspIleLysSerGluAspTrpSer 290
 DB 604 GTCAGATTCGCGAGCTCATTTGCTCTGAGATATCAAAACGCGTGAAGTCT 657

RESULT 13
 LOCUS BQ273302 515 bp mRNA linear EST 07-MAY-2002
 DEFINITION saoz3e12.y1 Gm-c1054 glycine max cDNA clone SOYBEAN CLONE ID:


```
Db 69 GAAGATCAGGAGGAAACCAACGTCCTACTGATGTGATGTGCATCGGTGCAGT 128
|||::: ||| |||||
QY 143 SerArgTyrLeuAlaLysLysPheGlyAlaThrSerValGlyIleThrLeuSerProVal 162
|||::: ||| |||||
Db 129 TCTAGTACCTTGCAGAAAGAAATATGCGCTAGAGCTAAAGGTATCATCTTGAGTCTGTA 188
|||::: ||| |||||
QY 163 GluAlaGlnArgAlaAsnAlaLeuAlaIleValThrGlnGlyLeuAlaAspLysValSerPhe 182
|||::: ||| |||||
Db 189 CAAGCAGAGAGAGGCTCAAGCTTGTGCTGATGCTCAGGATTAAGGTATGATTAAGGTTCATTT 248
|||::: ||| |||||
QY 183 GluValAlaAspAlaLeuGlnGlnProPheSerAspGlyGlnPheAspLeuValTyrSer 202
|||::: ||| |||||
Db 249 CAAGTAGAGAGCGCTTCAATCAGCTTTTCCAGATGGCAATTGCACTTGGTTGGTCC 308
|||::: ||| |||||
QY 203 MetGlnSerGlyGluHisMetProAspLysAlaLysPheValGlyGluLeuAlaArgVal 222
|||::: ||| |||||
Db 309 ATGGAGAGTGAGAAACACATGCGCCGAAAGAAAGTTGTTGGCAATTAGCTCGAGCTG 368
|||::: ||| |||||
QY 223 AlaAlaProGlyAlaIleIleIleValThrTyrCysHisArgAspLeuGlyProAsp 242
|||::: ||| |||||
Db 369 GCAGCACCAGAGGACCAATCATCTTGTCTCATGTGTCACAGGAGGACCTTCCCTTCG 428
|||::: ||| |||||
QY 243 GluGlnSerLeuHisProTyrGlnAspLeuLeuLysLysIleCysAspAlaTyrTyr 262
|||::: ||| |||||
Db 429 GAGCAATCTCTGACTCCAGAGAGAGAAAGAGCTTAATAATAGATATGCAAGCCCTCTAT 488
|||::: ||| |||||
QY 263 LeuProAlaTyrCysSerThrSerAspTyrValLysLeuGlnSerLeuSerLeuGln 282
|||::: ||| |||||
Db 489 CTTCGGCTGTGTCTCCACTGCTGATTAATGTAAGTTACTCAATCAATCTCTTCAG 548
|||::: ||| |||||
QY 283 AspIleLysSerGluAspTyrSerArgPheValAlaProPheTyrProAlaValIleArg 302
|||::: ||| |||||
Db 549 GATATCAAGGAGAGAGCTGGCTGAGAAATGTGCTCCATTTTGGCCAGCAGTCATTAAG 608
|||::: ||| |||||
QY 303 SerAlaPheThrTyrPlys 308
|||::: ||| |||||
Db 609 TCAGCAGTCGACATGGA 626
|||::: ||| |||||
```

Search completed: March 19, 2003, 08:24:12
Job time : 1297 secs